Title: Enhancer AAVs for targeting spinal motor neurons and descending motor pathways in rodents and macaque

Authors: Emily Kussick[1], Nelson Johansen[1], Naz Taskin[1], Brooke Wynalda[1], Refugio Martinez[1], Erin L. Groce[1], Melissa Reding[1], Elizabeth Liang[1], Lyudmila Shulga[1], Cindy Huang[1], Tamara Casper[1], Michael Clark[1], Windy Ho[1], Yuan Gao[1], Cindy T.J. van Velthoven[1], Cassandra Sobieski[1], Rebecca Ferrer[1], Melissa R. Berg[2], Britni C. Curtis[2], Chris English[2], Jesse C. Day[2], Michal Fortuna[1], Nicholas Donadio[1], Dakota Newman[1], Shenqin Yao[1], Anish Bhaswanth Chakka[1], Jeff Goldy[1], Amy Torkelson[1], Junitta B. Guzman[1], Rushil Chakrabarty[1], Beagen Nguy[1], Nathan Guilford[1], Trangthanh H. Pham[1], Vonn Wright[1], Kara Ronellenfitch[1], Kathryn Gudsnuk[1], Bargavi Thyagarajan[1], Kimberly A. Smith[1], Nick Dee[1], Hongkui Zeng[1,3], Zizhen Yao[1], Bosiljka Tasic[1], Boaz P. Levi[1], Rebecca Hodge[1],Trygve E. Bakken[1], Ed S. Lein[1,3], Jonathan T. Ting[1,3], Tanya L. Daigle[1,3,4*]

Author Affiliations and Footnotes:

[1] Allen Institute for Brain Science, Seattle, WA 98109, USA.
[2] Washington National Primate Research Center, Seattle, WA 98195 USA.
[3] University of Washington, Seattle, WA 98195 USA.
[4]Lead Contact
*Corresponding Author

Correspondence: Tanya L. Daigle, tanyad@alleninstitute.org

Summary:

Experimental access to cell types within the mammalian spinal cord is severely limited by the availability of genetic tools. To enable access to lower motor neurons (LMNs) and LMN subtypes, which function to integrate information from the brain and control movement through direct innervation of effector muscles, we generated single cell multiome datasets from mouse and macaque spinal cords and discovered putative enhancers for each neuronal population. We cloned these enhancers into adeno-associated viral vectors (AAVs) driving a reporter fluorophore and functionally screened them in mouse. The most promising candidate enhancers were then extensively characterized using imaging and molecular techniques and further tested in rat and macaque to show conservation of LMN labeling. Additionally, we combined enhancer elements into a single vector to achieve simultaneous labeling of upper motor neurons (UMNs) and LMNs. This unprecedented LMN toolkit will enable future investigations of cell type function across species and potential therapeutic interventions for human neurodegenerative diseases.

Keywords: AAV; enhancer; multiome; RNA-seq; ATAC-seq; spinal cord; motor neuron; cell types; macaque; rodents

Introduction

The spinal cord is a vital part of the central nervous system (CNS) and functions to relay information back and forth between the brain and the body. It is comprised of a diversity of cell

types that have been classically defined by their morphological and physiological properties, and more recently by their single cell transcriptomes [1–6]. Lower motor neurons (LMNs) are specialized types of cholinergic cells that reside in the brainstem and the ventral horn of the spinal cord and receive input from upper motor neurons (UMNs) located in the brain directly in humans or indirectly via an intermediate synapse in rodents and in turn, initiate actions by controlling effector muscles in the periphery [7]. Degeneration of motor neurons (MNs) can profoundly impact the initiation and control of movement and is a hallmark of many devastating diseases such as Amyotrophic Lateral Sclerosis (ALS) and Spinal Muscular Atrophy (SMA) for which there are no known cures [8,9].

While much progress has been made on LMN cell type definitions, there is still much to learn about their overall organization and interconnectivity within the spinal cord and with cells throughout the body in both healthy and pathological conditions. Experimental access to LMNs in rodents still largely depends on the Chat-IRES-Cre line which broadly labels all cholinergic neurons throughout the body [10] or vectors containing the homeobox gene *Hb9* derived promoter fragments [11–14] which exhibit varying levels of specificity and strength depending on the experimental context. Importantly, no genetic tools exist to selectively target LMNs across species or to target any LMN subtypes (i.e. alpha, beta, or gamma) in any species. This precludes the possibility of refined targeting strategies, such as those needed to better define cell type function and for potential therapeutic interventions.

We previously developed an approach to identify functional enhancer elements within the genome and used it to create adeno associated viruses (AAVs) to target specific interneuron and excitatory projection neuron subtypes within the mouse neocortex and to a limited degree in the non-human primate brain [15,16]. More recently, we scaled and expanded the effort to create larger suites of viral genetic tools for near complete coverage of the neocortical neuron landscape and of major non-neuronal types in mouse [17,18]. These successful efforts showed generalizability of the enhancer AAV technology platform and suggested the potential for future tool development outside of the brain.

Here, we provide high-quality, single cell multiome datasets from the mouse and macaque spinal cord to reveal genome-wide regions of open chromatin across major cell classes. We use these datasets to discover putative enhancer elements for LMNs and LMN subtypes with potential for cross-species conservation, and functionally test all of them in mouse and a subset of the most promising in rat and macaque using clinically relevant routes of administration to show they effectively target the homologous cell populations. Additionally, we demonstrate that enhancers targeting both UMN and LMN populations can be stitched in *cis* in a single viral vector to effectively achieve the combined labeling pattern within the CNS of rodents.

Results

Generation of 10X Multiome data from mouse and macaque spinal cord and identification of major cell classes

To identify LMN-specific enhancers, we isolated nuclei from cervical and lumbar segments of spinal cord in mouse and macaque. We selectively enriched for neurons using fluorescence-activated nuclei sorting (FANS) to isolate nuclei at a ratio of 70% NeuN+ and 20% OLIG2-, 10% OLIG2+ (see **Methods Details**). Gene expression and open chromatin were profiled from individual nuclei using 10X Multiome. In total, we collected high quality data from 73,120 single nuclei from 5 mice and 34,344 single nuclei from 4 macaques (**Figure 1A**). High quality nuclei

were retained that expressed more than 1000 genes, had less than 3% mitochondrial reads and doublet scores less than 0.3, and had ATAC-seq fragments in transcription start sites and that spanned nucleosomes based on standards reported by ENCODE [19].

Spinal cord cell types were defined based on iterative clustering of transcriptomic data, and all types had distinct marker genes (see **Methods Details**; [1,2]). Cell type identities were assigned based on enriched expression of known marker genes for LMNs (CHAT, ISL1, PRPH), other excitatory neurons (SLC17A7), inhibitory neurons (GAD1), and non-neuronal types (SLC1A3, MOG) (**Figure 1B**). Macaque and mouse snRNA-seq data was aligned using scVI [20]. We found consistent coverage of a conserved set of cell types across species (**Figure 1C**) highlighting conservation of distinct marker genes for spinal cord cell types.

Identification of putative pan motor neuron enhancers and deeper analysis of the 1137m enhancer

To identify putative enhancers specific to LMNs, we aggregated the snATAC-seq data for all nuclei within each species and cell class and identified peaks specific to mouse LMNs that were proximal to published LMN marker genes [1,2] (**Figure 1D**). Many pan-LMN enhancers had conserved accessibility between mouse and macaque (**Figure 1**). We further evaluated cell type- and tissue-specificity of each pan-LMN enhancer in human non-brain tissues [21] and in the whole mouse brain [22] by assessing the accessibility of the corresponding genomic region (**Table S1)**. 1137m was found to be highly specific to LMNs and inaccessible in all non-brain tissues (see **Methods Details**; **Figure 1E**). Additionally, based on mouse whole brain ATAC profiling, we found accessibility was limited to three rare cell types (HB Calcb Chol, POR Spp1 Gly-Gaba and STR-PAL Chst9 Gaba) (see region definitions in **Figure legend 1F**) that are predominantly found in brainstem (**Figure 1F; Table S1**). Interestingly, only HB Calcb Chol expresses the enhancer proximal LMN marker *Chat*, suggesting that 1137m has altered regulatory function in the other two cell types.

Generation of enhancer-driven fluorophore AAVs for pan spinal motor neuron labeling

We first assessed the ability of reported blood-brain-barrier (BBB) permeant capsids, PHP.eB and 9P31 [23,24] to effectively cross the blood-spinal cord barrier (BSCB) and transduce neurons broadly within the spinal cord following retro-orbital (RO) delivery to mice. We packaged recombinant AAVs (rAAVs) containing a pan neuronal human synapsin 1 (hSyn1) promoter driving a histone 2B (H2B) tethered directly to a SYFP2 fluorescent protein for nuclear enrichment called CN1839 [15] and delivered 5.0 x 10^{11} GC of PHP.eB or 9P31-serotyped CN1839 viruses to wild-type mice and assessed viral expression in brain and spinal cord 3-4 weeks post-infection (**Figure S1A, B**). We found that both serotypes exhibited strong nuclear SYFP2 neuronal labeling throughout the brain and in all levels of spinal cord, uniformly in dorsal and ventral horn regions (**Figure S1A, B**). For comparison, we also tested PHP.S serotyped virus, a broad peripheral nervous system (PNS) directed capsid with some reported spinal cord expression [23], and found the same GC gave very little expression in brain or spinal cord following RO administration (**Figure S1A, B**). These data suggested that either PHP.eB or 9P31 capsid would work for non-invasive screening of putative enhancers targeting spinal cord cell types, and therefore we selected the PHP.eB capsid because we have used it extensively in our previous work [15,16].

To functionally test the putative enhancers for LMNs, we cloned them into an rAAV genome upstream of a minimal β-globin promoter driving the SYFP2 fluorescent protein, delivered 5.0 x 10^{11} GC of PHP.eB viruses RO to wild-type mice, and analyzed expression in spinal cord 3-4 weeks post-infection (**Figure 2A**). We screened a total of 21 putative LMNs and found that 16 (76%) of these exhibited SYFP2+ cell body labeling within the transverse spinal cord section, 13 of these showed SYFP2+ labeling of cells within the ventral horn (61%) and 5 showed no SYFP2+ cells (24% failure rate) in the sections analyzed (**Figure S2 A, B**). The HCT1-1137m-SYFP2 enhancer virus (referred to as HCT1 from here forward) appeared initially to exhibit the strongest labeling of spinal motor neurons (**Figure 2B, and Figure S2A**) and therefore, we selected it for additional examination using a variety of experimental techniques. We performed light sheet imaging of cleared spinal cords from HCT1 virus infected animals and found native SYFP2+ cells within the ventral horn at all levels of spinal cord (**Figure 2C and Video S1**). To confirm the labeled cells were positive for choline acetyltransferase (Chat), we injected HCT1 virus into Chat-IRES-Cre;Ai75 mice double transgenic mice (**Figure 2D**) and evaluated co-localization of native SYFP2 and with nuclear tdTomato fluorescence throughout spinal cord (cervical, thoracic, lumbar, and sacral) and for this analysis, each spinal cord level was further subdivided coarsely into dorsal, medial, and ventral parts (**Figure 2E, F**). This analysis revealed that the HCT1 virus labeled LMNs with high specificity in ventral horn (83.8% cervical, 86.3% thoracic, 78.8% lumbar, and 93.2% sacral) (**Figure 2G**). It retained specificity in dorsal and medial regions, with values that ranged from 70.5% (lumbar medial) to 93.4% (sacral dorsal) and revealed a significant difference in the specificity achieved in dorsal lumbar vs. dorsal sacral segments (p=0.0253) (**Figure 2G**). The completeness of labeling for the HCT1 virus ranged from 50.4% (sacral medial) to 81.7% (thoracic dorsal) in the regions analyzed with significant differences observed between regions or subdivisions (**Figure 2H**). These results show that the HCT1 virus strongly labels LMNs, and cholinergic interneurons mostly enriched in medial and dorsal regions, which is consistent with the proximity of the 1137m enhancer to the *Chat* gene within the mouse genome.

To determine if the HCT1 virus could label LMNs in rats, a complimentary rodent species commonly used for biological research, we delivered ~1.50-2.00 X 10^{11} GC of it to rat neonatal pups via intracerebroventricular (ICV) injection and analyzed expression in cervical spinal cord 3-4 weeks post-infection (**Figure 3A**). Like in mouse, we found that the HCT1 virus strongly labeled LMNs within the ventral horn as validated by the co-localization of native SYFP2 fluorescence with anti-Chat immunolabeling (**Figure 3A, C**). The specificity and completeness of HCT1 virus labeling was found to be 86.6% and 68.5% respectively (**Figure3 D, E**). In contrast to that observed in mouse, we did not find significant labeling of Chat+ interneurons in cervical spinal cord sections. These data may suggest divergent enhancer activity within Chat+ interneurons or potentially a methodological difference between transgenic versus direct immunological approaches.

We next tested whether the HCT1-1137m enhancer would exhibit similar activity in LMNs in the non-human primate (NHP) spinal cord following two different clinically relevant routes of administration. We first constructed a vector containing the 1137m enhancer driving a 3XFlag epitope tagged mTFP1 fluorophore (referred to as HCT69 from here forward) and packaged into purified PHP.eB virus (**Figure 4A**). Naïve juvenile macaques were then injected with 1.0 X 10^{13} vg/kg of HCT69 virus via the intra-cisterna magna (ICM) or intrathecal (IT) routes of administration and sacrificed 5-7 weeks post-injection for histological analyses (**Figure 4A**; also see **Methods Details**). Fixed transverse sections from each animal, each level of spinal cord,

were evaluated for native mTFP1 fluorescence (data not shown) and stained with anti-Flag epitope and anti-Chat antibodies to reveal viral labeling and cholinergic positive LMNs and interneurons (**Figure 4B, C**). With both routes of administration, we observed highly specific viral labeling of LMNs in all levels of spinal cord (86.0-99.0% range; **Figure 4B-C**). We found significant differences in specificity between sacral and all other cord levels for IT route (IT; cervical vs sacral, p=0.0274; thoracic vs sacral, p=0.0008; lumbar vs sacral, p=0.0185) and in lumbar versus sacral cord for ICM route (p=0.0066) (**Figure 4D**). The completeness of labeling for the ICM route was similar across all levels of spinal cord ranging from 30.1-43.8% and for the IT route was unexpectedly highest in cervical and lumbar levels of cord and significantly lower in other regions (IT; cervical vs sacral, p=0.0001; lumbar vs sacral, p=0.0212) (**Figure 4E**).

We further evaluated HCT69 virus specificity in antibody-stained fixed sections prepared from the brain and peripheral tissue (heart, liver, kidney, muscle, and pancreas) of ICM and IT injected animals (**Figure S3B, C**). We additionally compared labeling to that achieved in an animal that received CN1839, a pan neuronal nuclear SYFP2-expressing virus, PHP.eB serotyped and delivered at the same dose via the ICM route of administration (**Figure S3A**). We found that the HCT69 virus showed very little expression in the brain (e.g. often none or a single FLAG+ cell within a given section) and no expression in any peripheral organ with either delivery route (**Figure S3B, C**). By comparison, the CN1839 virus gave moderate neuronal labeling throughout the brain, with notable pockets of expression in cortical regions (**Figure S3A**), consistent with previous reports of AAV9 or AAV9 based virus infectivity [25,26]. Additionally, we found widespread neuronal labeling in the cervical spinal cord of the CN1839 animal, but few SYFP2+/Chat+ LMNs were labeled in ventral horn (**Figure S3A**). In the periphery, we observed little (heart) to no expression (all other organs) (**Figure S3A**). A summary of the biodistribution data for each virus tested in macaque can be found in **Table S2**. These data collectively demonstrate that the HCT69 enhancer virus labels LMNs within the body more selectively than a hSyn promoter driven virus and are consistent with the predicted body-wide accessibility profile for the 1137m enhancer in human scATACseq data (**Figure 1D**).

Generation of enhancer-driven fluorophore AAVs for LMN subtypes

Somatic LMNs can be further divided into alpha, gamma and beta types depending on the type of muscle fiber innervated [6] and have been shown to be differentially susceptible in various spinal cord diseases [9,27,28] **(Figure 5A**). Therefore, we next sought to develop enhancer AAVs for alpha and gamma LMN subtypes for more refined targeting approaches. We discovered differentially accessible putative enhancers proximal to known marker genes, *Chodl* and *Ret,* for each subtype within the mouse and NHP spinal cord multiome datasets (**Figure 5B**) and proximal to other previously described marker genes (**Figure S4B)** [1,2,29]. We cloned 20 putative alpha LMN subtype and 3 putative gamma LMN subtype enhancers into the SYFP2-expressing rAAV backbone and screened PHP.eB viruses in wild-type mice following RO injection (**Figure 5C, D** and **Figure S4A, B**). We found 16 out of 20 alpha subtype LMN enhancers (80%) showed SYFP2+ labeling of very large cell bodies within transverse sections from cervical spinal cord, which is consistent with the described morphology, and of these 20 enhancers, 10 labeled cells within the ventral horn (50%) (**Figure 5D and Figure S4A, B**). For putative gamma subtype LMN enhancers, we found that 3 out of the 3 (100%) showed SYFP2+ labeling of very small cell bodies within the ventral horn in transverse cervical cord sections (**Figure 5D and Figure S4A, B**). The HCT46-1188m-SYFP2 virus (referred to as HCT46 from here forward)

exhibited the strongest and most abundant labeling of potential gamma subtype LMNs and therefore this virus was selected for further characterization. The HCT3-1139m-SYFP2 virus initially exhibited very specific labeling of potential alpha subtype LMNs but appeared somewhat weak. Therefore we optimized it by concatenation of the putative enhancer core as previously described [16], and created a derivative virus termed HCT49-3Xcore-1139m-SYFP2 (referred to as HCT49 from here forward; **Figure 5C**) for further characterization.

To validate HCT46 and HCT49 viruses, we evaluated expression in Chat-IRES-Cre;Ai75 double transgenic mice following RO injection of each virus. We found co-localization of SYFP2+ cells with Chat+ cells in the ventral horn of transverse cervical spinal cord sections, confirming labeling of LMNs (**Figure 5E**). To validate each virus at the level of subtype, we performed RNAscope analysis on fresh frozen cervical spinal cord sections using a combination of *SYFP2*, *Chat*, and *Ret* or *Chodl* probes for the HCT46 or the HCT49 viruses, respectively (**see Methods Details**; **Figure 5F, G**). Here we observed the expected co-labeling of *Chat* signal with *Ret* or *Chodl* signals in single cells within HCT46 virus-infected sections (**Figure 5F, left panels**) or HCT49-virus infected sections (**Figure 5G, left panels**), which confirmed subtype probe specificity. We then evaluated the specificity of the HCT46 or HCT49 viruses and found a high degree of overlap between *SYFP2*+ and *Ret*+ or *Chodl*+ cells in cervical spinal cord sections (**Figure 5F, G, right panels**). Analysis of these data revealed that both HCT46 and HCT49 viruses exhibit high specificity for gamma and alpha subtype LMNs (85.8% and 92.5%, respectively; **Figure 5H, I, left panels**). Additionally, we found that more than half of alpha subtype LMNs were labeled by the HCT49 virus (58.1%; **Figure 5I, right panel**) and 39.1% of gamma subtype LMNs were labeled by the HCT46 virus in cervical spinal cord sections (**Figure 5H, right panel**).

Enhancer stitching enables dual labeling of UMNs and LMNs by a single virus

We have previously shown that enhancer viruses targeting different cell types can be co-delivered to achieve simultaneous labeling in one animal [16]. To extend this approach and simplify it, we next sought to develop a single vector system to achieve combinatorial labeling of cell types within the brain and spinal cord. We targeted UMNs of the corticospinal tract with LMNs (**Figure 6A**) because these cell types selectively degenerate in diseases such as Amyotrophic Lateral Sclerosis (ALS) [27,30] suggesting potential therapeutic applications for such a vector. Additionally, this approach was pursued because we have a large collection of well validated enhancers for each cell type UMNs and LMNs in hand [16] and present study). We cloned the UMN enhancer 453m upstream of the LMN 1137m enhancer and the minimal beta-globin promoter in a SYFP2-expressing viral vector (**Figure 6B**). We then packaged HCT55-453m-1137m-SYFP2 (referred to as HCT55 from here on) and 453m or 1137m single enhancer control vectors into PHP.eB serotyped viruses and performed RO injections in wild-type mice (**Figure 6C**). As expected, the 453m enhancer alone drove robust native SYFP2 fluorescence of UMNs within layer 5 of the cortex throughout the full rostral caudal axis, albeit strongest in posterior cortex, and in neurons within the deep cerebellar nucleus (DCN) and within lamina 10 (10Sp) of the spinal cord (**Figure 6C, left panels**). The 1137m enhancer alone drove robust SYFP2 expression in LMNs and in several hindbrain structures (**Figure 6C, middle panels**) as predicted from the mouse whole brain ATAC profiling analysis (**Figure 1F**). With the HCT55 stitched virus, we found the combined robust native SYFP2 expression in UMNs throughout the cortex and in neurons enriched in the ventral horn of the spinal cord (**Figure 6C, right panels**), some of which were confirmed to be LMNs following additional validation in Chat-IRES-Cre;Ai75

double transgenic mice (**Figure 6D**). Analysis of SYFP2+ and tdTomato+ cells revealed that the HCT55 virus labeled 40.8% of LMNs in cervical cord and was found to be 60.3% specific (**Figure 6E**), which is lower for both measures compared to HCT1 viral labeling in spinal cord (**Figure 2G, H**).

To determine if a different route of administration would improve HCT55 viral labeling of both UMNs and LMNs, we performed ICV injections of mouse neonatal pups and evaluated native SYFP2 fluorescence in the brain and spinal cord several weeks post-injection (**Figure 6F**). Here we found extremely robust labeling of UMNs in layer 5 throughout the cortex and a notable decrease in the amount of L2/3 pyramidal neuron expression (**Figure 6F, bottom panels**). Additionally, we observed more selective labeling of LMNs within the ventral horn in cervical spinal cord sections (**Figure 6F, top panels**). Taken together, these data suggest that the ICV route of administration may improve HCT55 virus specificity in both brain and the spinal cord.

Lastly, we evaluated the ability of the HCT55 virus to label both UMNs and LMNs across rodent species. The virus was injected into rat neonatal pups via the ICV route of administration and native SYFP2 fluorescence was evaluated in the target tissues several weeks post-injection (**Figure 6G**). In these experiments, we observed robust and specific SYFP2 labeling of LMNs in cervical spinal cord (**Figure 6G, top panels**) and labeling of UMNs in cortical layer 5 with most prominent expression within caudal regions of the brain (**Figure 6G, bottom panels**). These data demonstrate that the individual enhancer activity is conserved between rodents in the stitched vector design.

Discussion

Currently available single cell ATAC-seq data sets for analyzing chromatin accessibility in the rodent spinal cord are limited to very early in development [31–33] and are non-existent in the case of primate spinal cord. Therefore, the 10X multiome datasets, which reveal both the transcriptome and epigenome in the same cell, generated in the present study from adult mouse and adolescent macaque represent a unique and tremendously valuable resource to the community. We have provided a roadmap for how these datasets can be used to identify functional enhancers and created a suite of MN selective viral tools for labeling of cell types in the adult rodent and NHP spinal cord. With additional analyses, these multiome datasets may enable enhancer discovery and tool development beyond MNs. This has the potential to re-define genetic approaches for targeting a diversity spinal cord cell types and overall cell type definition across species.

Our previous efforts established the enhancer AAV technology platform as a robust approach for brain cell type-specific AAV tool development [15–18]. In the present work, we used similar methodology for one at a time screening and validation of candidate enhancers for pan MNs and MN subtypes and had a high success rate for translating enhancer discovery into a functional tool (> 75%) with equivalent or greater than we had previously targeting brain cell types. These results, combined with our success developing enhancer AAVs for targeting cell types outside of the CNS (data not shown), suggest that any cell type within the body is within reach. Future work in this space will rely heavily on the availability of high quality epigenomic data with sufficient cell type resolution such as reported in [21,34] and optimization of other experimental parameters (i.e. capsid, delivery route, and dose) to ensure success.

The single 1137m enhancer-containing viruses (HCT1 and HCT69) provided highly specific labeling of LMNs in mouse, rat, and macaque spinal cord following different routes of

administration. Variations in the extent of labeling (i.e., completeness) depended on the route of administration as we have seen previously with enhancer AAVs targeting brain cell types [16]. Similarly, the administration route used influenced the extent of labeling observed between levels of spinal cord analyzed. We unexpectedly found that the IT route of administration in macaque yielded the most complete viral labeling in cervical cord rather than in lumbar cord, the latter being the level at which the virus was administered into the animal. This may reflect a technical issue (e.g. missed injection site) or potentially an unknown biological mechanism not fully appreciated at this time, and that we cannot fully resolve without an additional replicate. We think it's likely that the completeness of labeling in macaque spinal cord could be improved with delivery of a higher dose of virus by either route of administration. These viruses represent useful research tools for direct analysis of LMNs in rodents and macaque and can be used alone or in combination with existing genetic tools, and/or could be further modified to drive functional transgenes as previously shown for other enhancer AAVs [15–17].

For potential translational applications, such as for human gene therapy, we evaluated 1137m enhancer-driven expression in the macaque brain and peripheral organs following moderately invasive relative to intraparenchymal routes of administration and found essentially no expression outside of LMNs within the spinal cord and this was notably more specific than that observed with a hSyn1 promoter-driven transgene. A lack of off-target viral expression in vital organs such as the heart and liver are desirable for potential gene therapies, as it is expression there and in dorsal root ganglion (DRGs) that are a consequence of ubiquitous promoter (pan cellular or pan neuronal like hSyn1) use and often underlie the clinical failures [35]. Therefore, the 1137m enhancer has the ideal body-wide biodistribution profile, which suggests that it could be deployed in therapeutic strategies and delivered less invasively as a PHP.eB serotyped virus via ICM or IT routes to access LMNs throughout the entire primate spinal cord.

Differential vulnerability of LMN subtypes in neurodegenerative diseases is a recurring theme and therefore, we generated enhancer AAVs targeting alpha or gamma type LMNs and showed high selectively of these viruses for *Chodl*+ fast-firing (FF) alpha type LMNs and *Ret*+ gamma type LMNs respectively in mouse spinal cord. Much less is known about *Chodl*+ FF alpha type LMNs and *Ret*+ gamma subtype LMNs and other transcriptomically-defined LMN subtypes due to their recent discovery [2] and the lack of available genetic tools to selectively access these cell types. It is generally accepted that at the broader subclass level alpha type LMNs selectively degenerate in diseases such as ALS, while gamma type LMNs are spared [8,36]. Therefore, subtype-selective LMN interventions enabled by our enhancer AAVs may be beneficial to slow disease progression and/or to provide support to the spared neuronal population. Accessibility of the functionally characterized subtype selective LMN enhancers were found to be conserved in the macaque spinal cord multiome dataset and therefore they may be portable for use in primate.

The ability to target multiple cell types selectively and simultaneously in a flexible manner using a single enhancer-based viral vector offers many advantages and increases the potential for use in therapeutic approaches. As proof of concept for this general approach, and for use in the ALS disease context, we stitched validated enhancers for UMNs and LMNs together to target descending motor pathways and found that the combined labeling pattern could be achieved in the CNS, although the specificity and extent of labeling was not as high as that obtained with the single enhancer viruses alone following RO delivery, indicating that improvements are needed to obtain the optimal design for these enhancers and for this route of administration.

However, it is notable that ICV delivery of the same virus to rodents appeared to improve the specificity, strength, extent of labeling in both the brain and the spinal cord and therefore may be an alternative approach to consider when greater specificity and completeness of labeling are needed. Given that both the individual UMN and LMN enhancers contained within the stitched design have been shown to label the homologous cell type in primate spinal cord and brain (as demonstrated in the current paper and additional data not shown), it seems likely that the combined enhancer vector would exhibit the same specificity. This combined enhancer vector could then be used with a recently discovered BBB-penetrant capsid (e.g. bCap1, Dyno Therapeutics unpublished) to achieve simultaneous targeting of a therapeutic cargo with a less invasive intravenous (IV) route of administration.

The suite of enhancer AAVs reported here will enable diverse cell type-specific applications, which will ultimately lead to a better understanding of MN function in both healthy and diseased states, and potentially more effective therapeutic interventions for debilitating and lethal neurodegenerative disorders.

Acknowledgements

We are grateful for the support of the following Allen Institute teams and departments: Transgenic Colony Management, Genotyping, Lab Animal Services, Viral Technology Core, Tissue Processing, Molecular Biology, Technology, and Human Cell Types. We thank Leah Chestnut and Brian Nguyen at LifeCanvas Technology for lightsheet imaging, video, and image preparation. We thank members of the Washington National Primate Research Center for providing animal care, surgical support, and assistance with tissue collection for the study. Schematics in many of the figures were created with BioRender.com. This work was supported by award number U19MH114830 from the National Institute of Mental Health to H.Z. and by BioMarin Pharmaceutical and the Paul G. Allen Foundation. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH, the National Institute of Mental Health or of BioMarin Pharmaceutical. We thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support.

Author contributions

T.L.D designed the study. R.H., C.S., and R.F. performed macaque spinal cord tissue collections, R.H., T.E.B., and N.J. oversaw 10X multiome generation from macaque, and performed analyses. N.D. managed tissue processing team and T.C., M.C., W.H., performed mouse spinal cord tissue collections for multiome data generation. Z.Y., Y.G., C.T.J.vV., B.T., N.J., and H.Z. oversaw 10X multiome generation from mouse and performed analyses. E.K, N.T., B.W., R.M., E.L.G., performed molecular cloning, histology, imaging, and analysis. N.T. M.R.B, B.C.C., C.E., J.C.D., and M.F. performed macaque injections and collected and processed macaque tissue. S.Y., N.D., and D.N. generated viruses. M.R., E.L, L.S., and C.H., performed mouse tissue collections. E.K. performed RNAscope experiments and analyses. J.T.T. oversaw rat experiments and protocol, and with E.K. performed tissue collection, imaging, and analyses. J.T.T. and K.G. managed macaque experiments and protocol. K.A.S. managed molecular biology team members A.B.C., J.G., A.T., J.B.G., R.C., B.N., N.G., T.H.P. and multiome data production. E.S.L. provided leadership of department and Gene Therapy

program. J.T.T., B.P.L., T.L.D. provided leadership of Gene Therapy and Viral Genetic Tools programs. T.L.D., N.J., E.K., and N.T. wrote the manuscript with support from all co-authors.

Declaration of interests

T.L.D., B.P.L, B.T., E.S.L, H.Z, J.T.T., N.J., Y.G., Z.Y., are inventors on one or more U.S. provisional patent applications related to this work. All authors declare no other competing interests.

Figures and Figure Legends

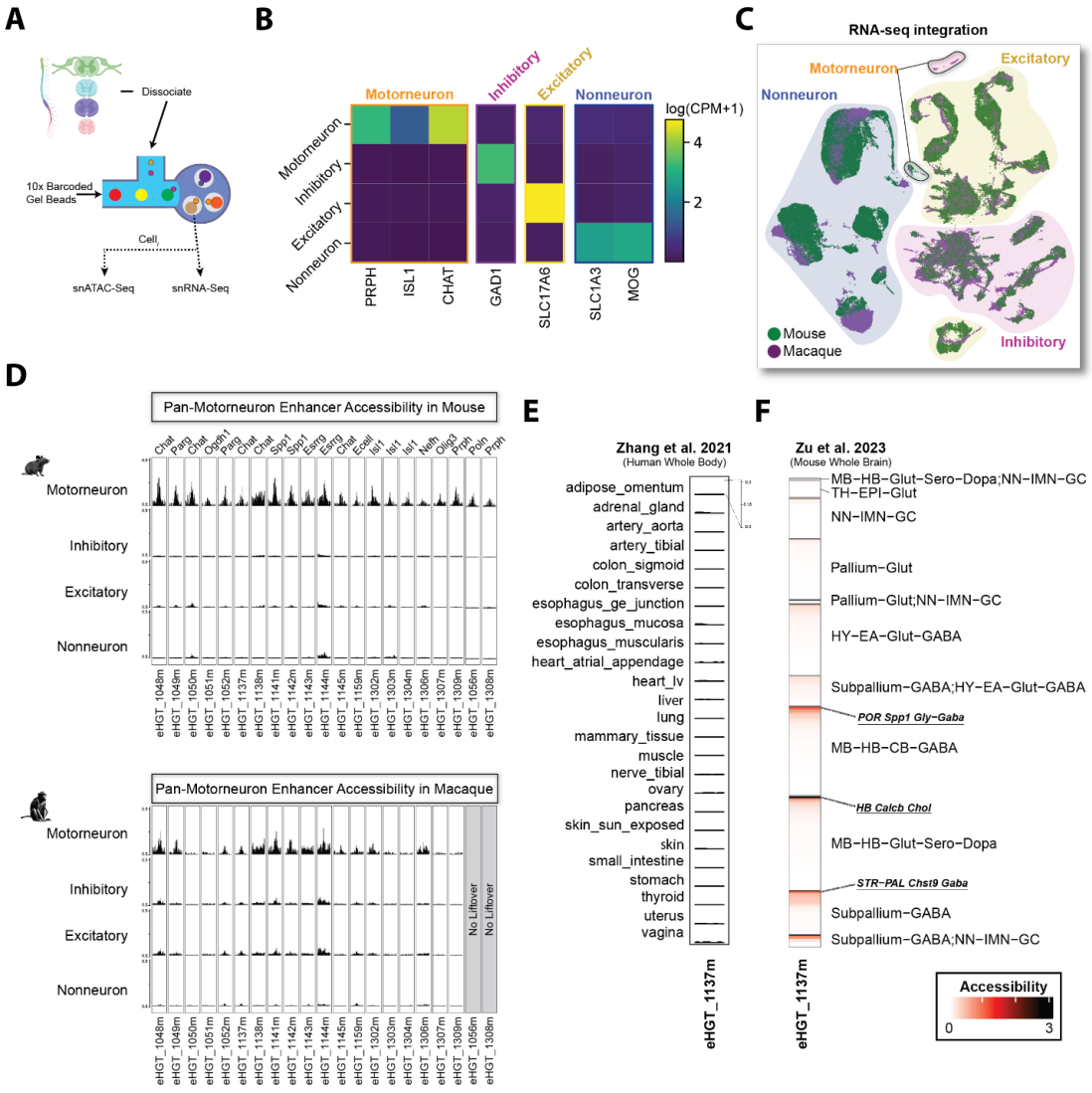


Figure 1. Generation of 10X multiome datasets from mouse and macaque spinal cord and motor neuron enhancer discovery. (**A**) Schematic of 10X Multiome profiling of spinal cord. (**B**) Heatmap of log2(CPM+1) expression of spinal cord marker genes. Colored boxes group the marker genes by broad spinal cord cell types. (**C**) Uniform manifold approximation and projection (UMAP) dimensional reduction of integrated snRNA-seq spinal cord data from Mouse

and Macaque. (**D**) Chromatin accessibility of pan-motorneuron enhancers in both Mouse and Macaque. (**E**) Chromatin accessibility of pan-motorneuron enhancer eHGT_1137m in multiple non-brain tissues across the human body (Zhang et al. 2021) (**F**) Chromatin accessibility of pan-motorneuron enhancer eHGT_1137m across the mouse whole brain (Li et al. 2023). Each row indicates the normalized accessibility of eHGT_1137m within each subclass (n=318) grouped by neighborhoods (n=11) from Yao et al. 2023. Subclasses in which significant expression was observed are highlighted in bold and underlined text. MB=midbrain, HB=hindbrain, TH=thalamus, EPI=epithalamus, HY=hypothalamus, CB=cerebellum.

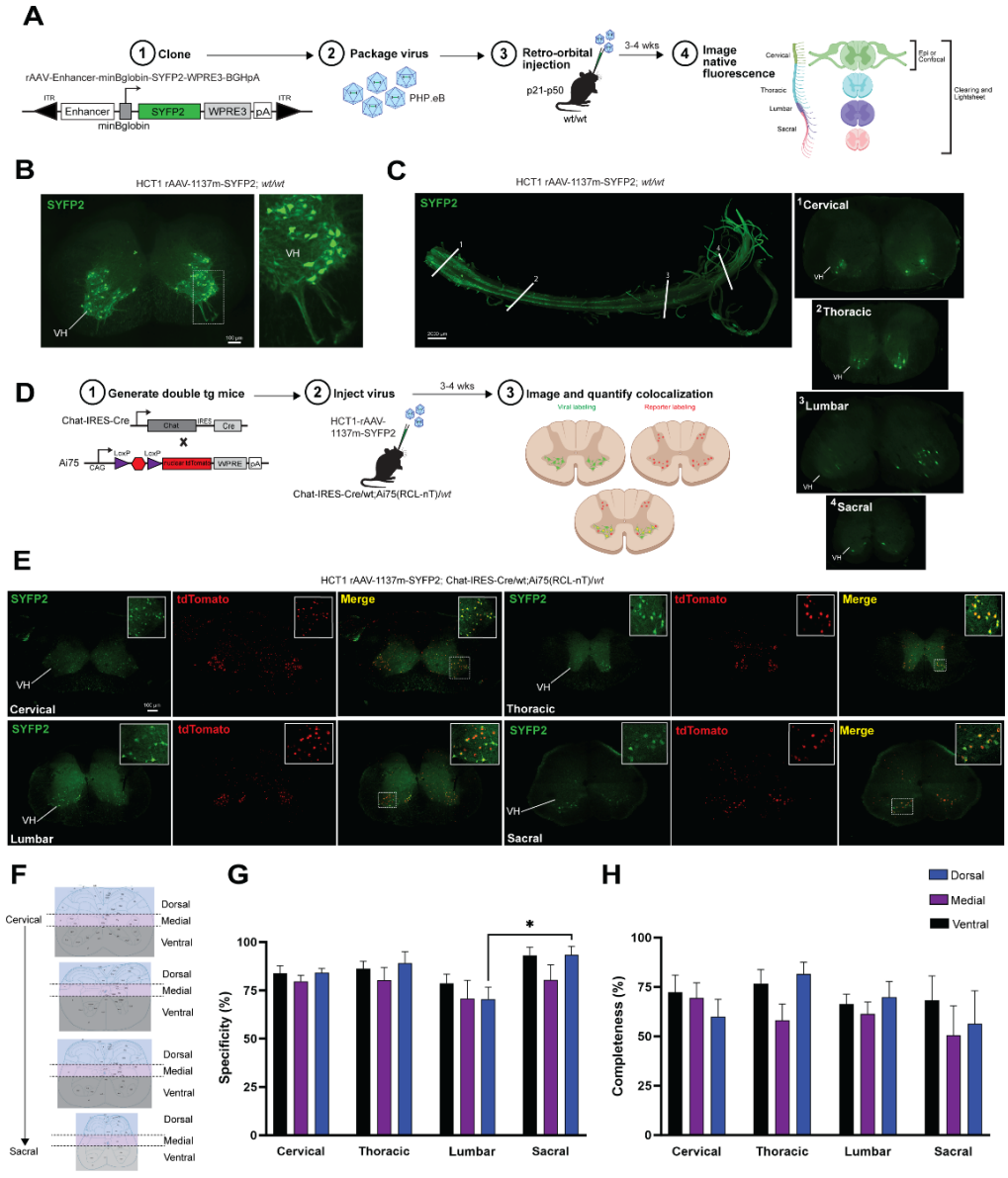


Figure 2. HCT1 enhancer AAV labels motor neurons in all levels of mouse spinal cord. (**A**) Experimental workflow for enhancer AAV generation and primary screening in mouse. Putative enhancers are cloned into a standard rAAV backbone containing the indicated components. minBglobin = minimal beta-globin promoter, SYFP2 = super yellow fluorescent protein 2,

WPRE3 = shortened woodchuck hepatitis virus posttranscriptional regulatory element, pA = bovine growth hormone poly A, ITR = inverted terminal repeats. 5.0 x 10^{11} GC of PHP.eB virus was delivered via retroorbital (RO) route of administration and spinal cord harvested 3–4-week post-injection and analyzed (**B**) Representative epifluorescence images of HCT1-driven native SYFP2 fluorescence in transverse spinal cord sections of 50µm thickness. VH = ventral horn and dashed box denotes area of interest highlighted in the right panel. (**C**) Lightsheet image of native SYFP2 fluorescence in intact spinal cord. Labels 1-4 denote the location of each 1.8µm transverse section shown in the right cascading panel. (**D**) Experimental workflow for validation of HCT1 virus using double transgenic mice. (**E**) Representative confocal images of native SYFP2 (green) fluorescence, nuclear tdTomato (red) fluorescence, and a merge of fluorescence from both channels (yellow) in transverse spinal cord sections from double transgenic mice. (**F**) Schematic of cell counting plan. Colored areas highlight coarse segmentation overlayed onto reference images from the Allen Institute mouse spinal cord atlas. (**G-H**) Quantification of HCT1 virus specificity and completeness of labeling throughout all segments and levels of cord. Data are plotted as mean ± SEM and *p < 0.05 (n = 2 animals and 3 sections for each level per animal analyzed).

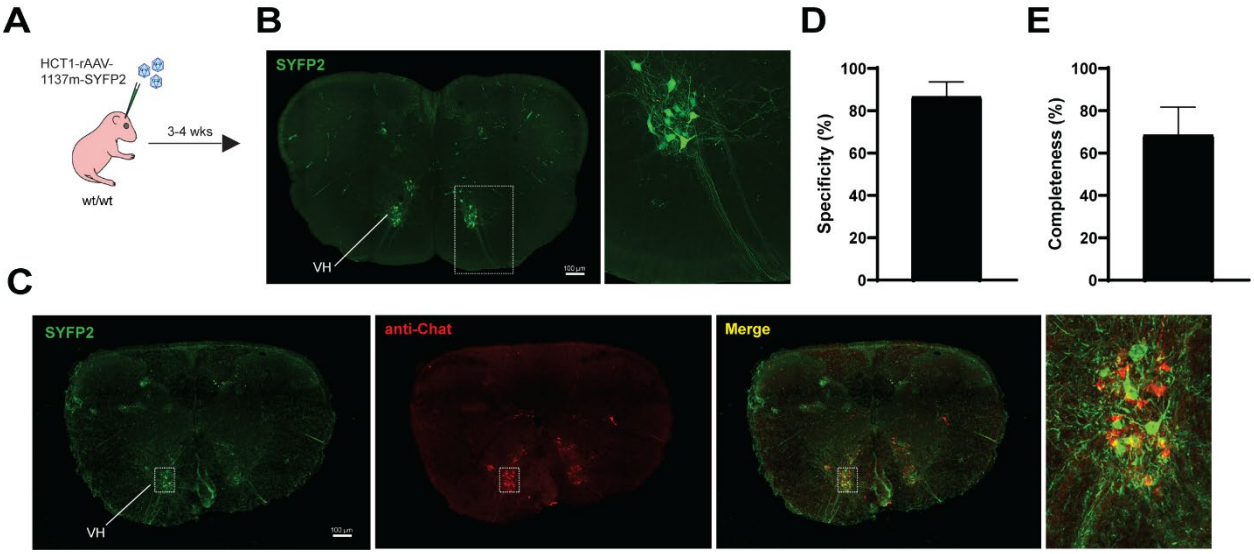


Figure 3. Conserved labeling of spinal motor neurons by HCT1 enhancer AAV in rat. (**A**) Schematic of HCT1 virus delivery via ICV route to rat neonatal pups. (**B**) Representative confocal images of HCT1-driven SYFP2 fluorescence in transverse cervical spinal cord sections of 50µm thickness. VH = ventral horn and dashed box denotes area of interest highlighted in the right panel. (**C**) Representative confocal images of native SYFP2 expression (green) and anti-ChAT immunostaining (red) in transverse cervical cord sections. (**D-E**) Quantification of HCT1 virus specificity and completeness of labeling in cervical cord. Data are plotted as mean ± SEM and n = 2 animals and 3 sections per animal analyzed.

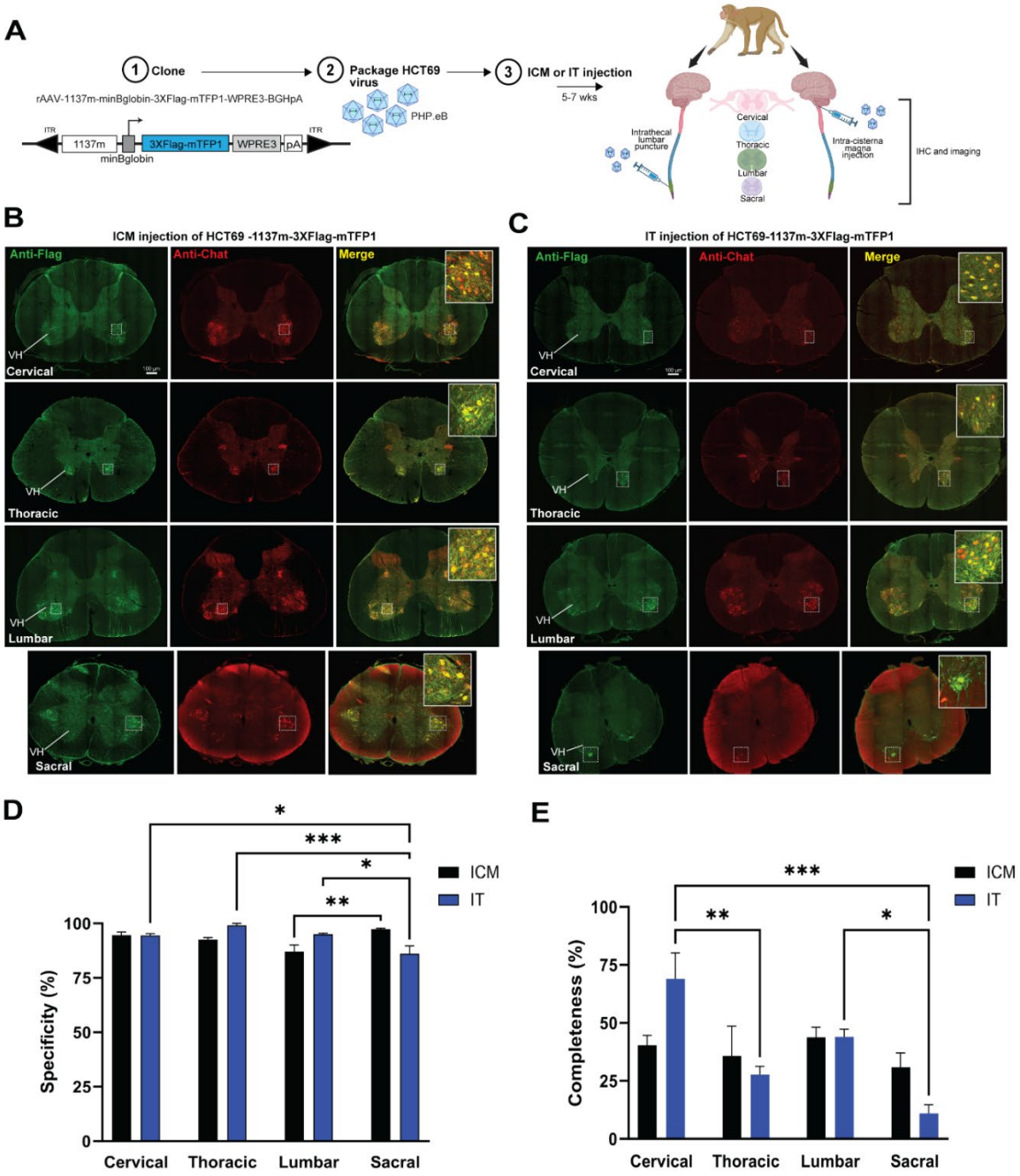


Figure 4. Conserved labeling of spinal motor neurons in macaque following clinically relevant routes of administration. (**A**) Experimental workflow for HCT69 enhancer AAV generation and testing in macaque by intra-cisternal magna (ICM) or intrathecal (IT) routes of administration. A 3XFlag-mTFP1 gene was inserted in place of SYFP2 in the standard rAAV backbone and packaged in PHP.eB serotype. Purified virus was delivered ICM or IT to young adult macaque animals and spinal cord and other organs were harvested 5-7 weeks post-injection. (**B-C**) Representative images of anti-Flag signal (red), anti-Chat signal (red) and a merge of both channels (yellow) in transverse spinal cord sections prepared from ICM (**B**) or IT (**C**) HCT69 injected animal. The level of spinal cord for each image is indicated. VH = ventral horn and dashed box denotes area of interest highlighted in the merge panel inset. (**D-E**)

Quantification of HCT69 virus specificity and completeness of labeling in all levels of macaque spinal cord. Data are plotted as mean ± S.D. and n = 1 animal per route and 3-5 sections per animal were analyzed. *p < 0.05, **p < 0.01, and ***p < 0.001 by one-way or two-way ANOVA.

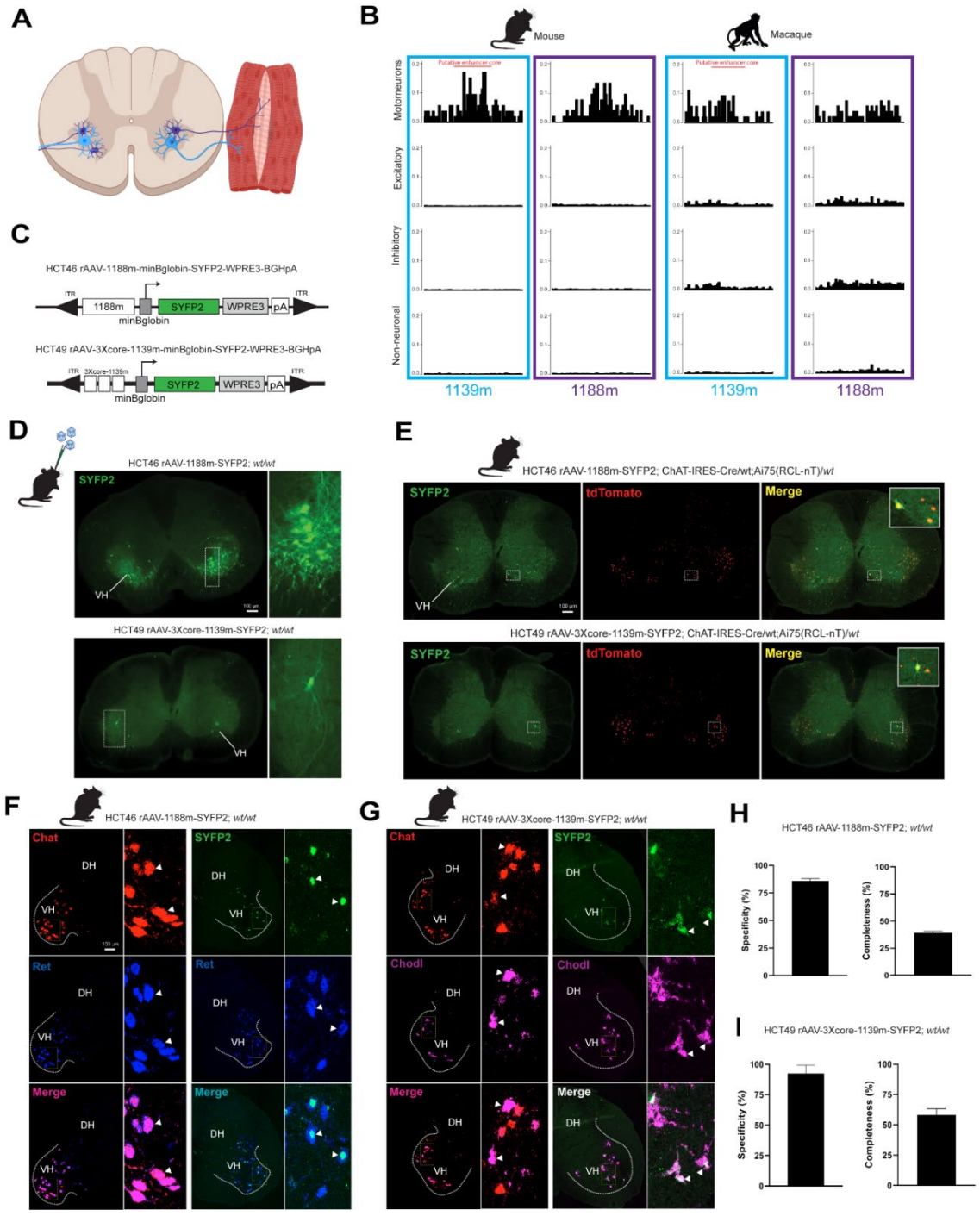


Figure 5. Enhancer AAVs for selective labeling of alpha or gamma spinal motor neurons in mouse cord. (**A**) Schematic of transverse cervical spinal cord section with alpha motor

neurons (blue) and gamma motor neurons (purple) residing in ventral horn and sending projections to extrafusal or intrafusal skeletal muscle fibers respectively. (**B**) Chromatin accessibility for 1139m and 1188m putative enhancers in mouse and macaque spinal cord. Each individual tract represents the indicated spinal cord cell type and the y-axis is normalized accessibility within the genomic region of each enhancer. (**C**) HCT46 and HCT49 vector diagrams. The 1188m enhancer or a 3Xcore of the 1139m enhancer was cloned into the standard rAAV backbone. (**D**) Representative epifluorescence images of HCT46- or HCT49-driven native SYFP2 fluorescence in transverse spinal cord sections of 50µm thickness. VH = ventral horn and dashed box denotes area of interest highlighted in the merge image inset. (**E**) Representative confocal images of native SYFP2 (green) fluorescence, nuclear tdTomato (red) fluorescence, and a merge of fluorescence from both channels (yellow) in transverse cervical spinal cord sections from double transgenic mice. Top panels are from a HCT46 injected animal and bottom panels are from a HCT49 injected animal. (**F-G**) Representative RNAscope images of hemisected transverse cervical cord sections from HCT46 (**F**) or HCT49 (**G**) virus injected animals. Probes for *Chat* (red), *Ret* (blue), *SYFP2* (green) and *Chodl (magenta)* were used in two-color experiments to confirm probe specificity and validate each virus. Curved dashed lines indicate the ventral horn (VH) and DH = dorsal horn; Dashed yellow box in hemisected images denote the region of interest in the panels to the right. White arrowheads indicate some of the double positive cells (e.g., *Chat*+/*Ret*+) for each condition. (**H-I**) Quantification of HCT46 (**H**) or HCT49 (**I**) virus specificity and completeness of labeling in cervical mouse spinal cord sections. Data are plotted as mean ± S.E.M and n = 2 per virus and 3-5 sections per animal were analyzed.

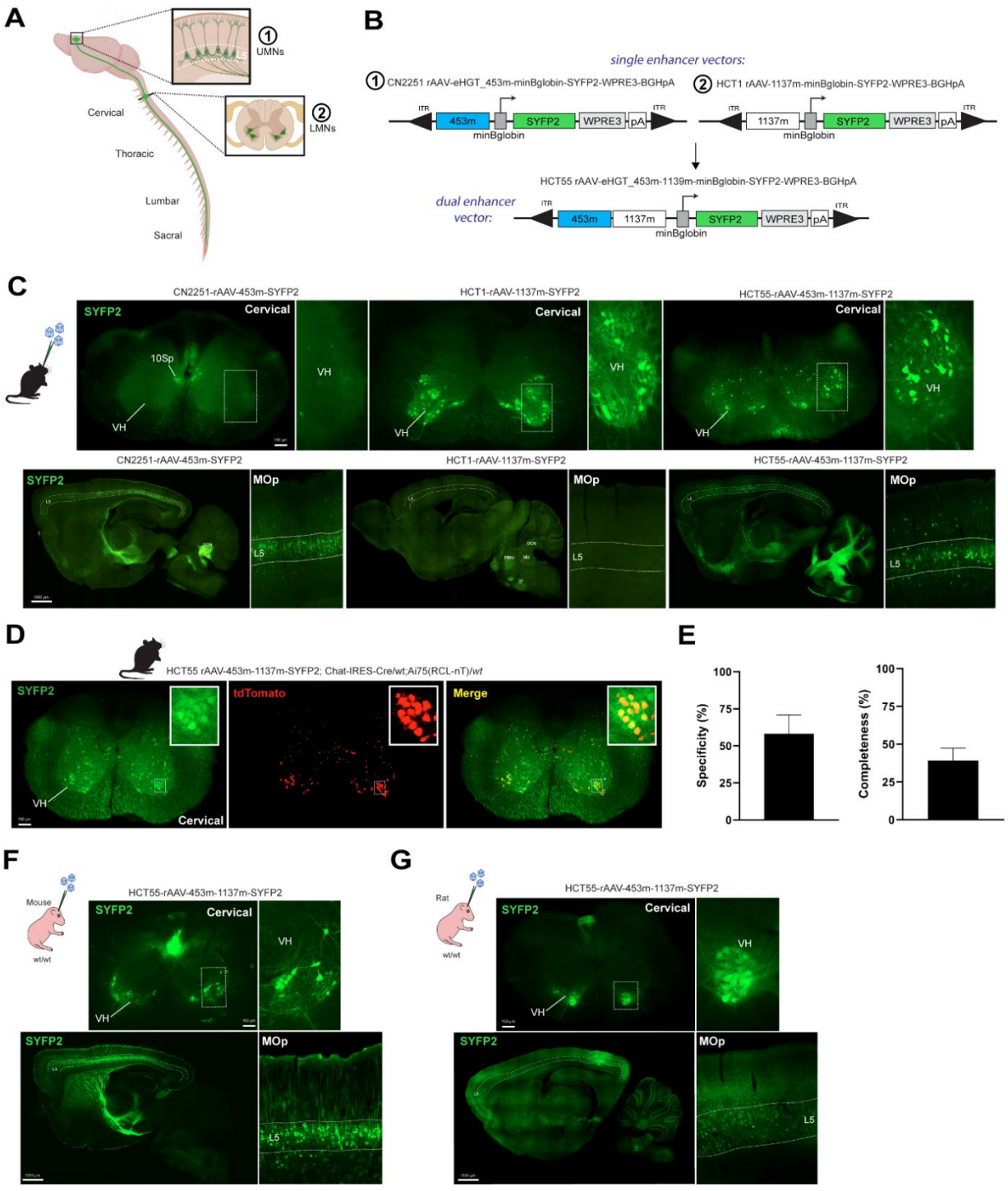


Figure 6. Enhancer stitching enables dual fluorophore labeling of upper and lower motor neurons in a signal viral vector. (**A**) Schematic of mouse central nervous system showing brain and spinal cord. Upper motor neurons (UMNs) located within layer 5 (L5) of the motor cortex (M1) are highlighted in the 3D brain structure and the right panel (1), and Lower motor neurons (LMNs) located in the ventral horn of the spinal cord are highlighted in a transverse cervical section in the right panel (2). The line traveling from brain to spinal cord represents the corticospinal tract. (**B**) Schematics of CN2251, HCT1, and HCT55 vector designs. In the single

enhancer vectors (CN2251 and HCT1), the 453m and 1137m enhancer sequences were cloned into the standard backbone upstream of the minBglobin promoter. In the dual enhancer vector (HCT55), the 453m enhancer was cloned directly upstream of the 1137m enhancer. (**C**) Representative epifluorescence images of virus driven SYFP2 expression in transverse spinal cord sections (upper panels) and in sagittal brain sections (lower panels). VH = ventral horn, 10Sp = lamina 10, DCN = deep cerebellar nucleus, PRNr = pontine reticular nucleus, MV = medial vestibular nucleus, VII facial motor nucleus, and the dashed boxes indicate the region of interest highlighted in the right panels. Dashed lines on the sagittal brain sections denote cortical L5 across the full rostral-caudal axis and representative images from M1 are shown in the right panels. (**D**) Representative confocal images of native SYFP2 (green) fluorescence, nuclear tdTomato (red) fluorescence, and a merge of fluorescence from both channels (yellow) in transverse cervical spinal cord sections from double transgenic mice injected with HCT55 virus. Insets show cell bodies in individual and merged channels. (**E**) Quantification of HCT55 virus specificity and completeness of labeling in double transgenic mice. Data are plotted as mean ± S.E.M and n = 2 animals and 3-5 sections per animal were analyzed. (**F-G**) Representative epifluorescence images of SYFP2 fluorescence in transverse cervical cord sections or brain sagittal sections following ICV delivery of HCT55 virus into mouse (**F**) or rat neonatal pups (**G**).

STAR Methods

Resource availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Tanya L. Daigle (tanyad@alleninstitute.org).

Materials availability

All plasmids generated in this study have been deposited to Addgene.

Data and Code availability

Newly generated 10x Multiome datasets from mouse and macaque have deposited to NeMO:

Mouse 10X Multiome snRNA-Seq
https://data.nemoarchive.org/biccn/grant/u19_zeng/zeng/multimodal/sncell/10xMultiome_RNAseq/mouse/raw/

Mouse 10X Multiome snATAC-Seq
https://data.nemoarchive.org/biccn/grant/u19_zeng/zeng/multimodal/sncell/10xMultiome_ATACseq/mouse/raw/

Macaque 10X Multiome snRNA-Seq
https://data.nemoarchive.org/other/grant/aibs_internal/lein/multimodal/sncell/10xMultiome_RNAseq/macaque/raw/

Macaque 10X Multiome snATAC-Seq
https://data.nemoarchive.org/other/grant/aibs_internal/lein/multimodal/sncell/10xMultiome_ATACseq/macaque/raw/

The Nemo identifier is nemo:dat-vccmgkw and can be found at https://assets.nemoarchive.org/dat-vccmgkw. Software code used for data analysis and visualization is available from GitHub at https://github.com/AllenInstitute/kussick2024analysis/. R packages for constructing and annotating cell type taxonomies can be found on GitHub at https://github.com/AllenInstitute/scrattch.taxonomy and https://github.com/AllenInstitute/scrattch.mapping respectively.

Experimental model and subject details

Mouse breeding and husbandry

Adult C57Bl/6J mice were bred and housed under Institutional Care and Use Committee protocols 2202 and 2301 at the Allen Institute for Brain Science, with no more than five animals per cage, maintained on a 12 day/night cycle, with food and water provided *ad libitum*. Juvenile C57BL/6J mice were also purchased directly from The Jackson Laboratory (Strain #000664) and housed at the Allen Institute for Brain Science. Both male and female mice were used for experiments and the minimal number of animals was used for each experimental group. Animals with anophthalmia or microphthalmia or other obvious conditions (e.g. penile prolapse) were excluded from experiments and all animals were maintained on a C57BL/6J background.

The Chat-IRES-Cre knock-in mouse line was purchased from The Jackson Laboratory (Strain #006410) and maintained on a C57BL/6J background. The *Rosa26* reporter line Ai75 (RCL-nls-tdTomato) line was generated in house [37] and maintained on a C57BL/6J background. Only mice heterozygous for both reporter and Cre driver were used for experiments.

Rat husbandry

Timed-pregnant female CD-1 Sprague-Dawley rats were purchased from Charles River Laboratories (Strain code #001) and utilized for experiments under Allen Institute Institutional Care and Use Committee protocol #2010. Animals were house no more than 1 per cage, maintained on a 12 day/night cycle, with food and water provided *ad libitum*. Rat pups were born generally the week following shipment receipt and at P1, the pups were tattooed for identification purposes and used for unilateral ICV injection of AAV vectors.

Macaque husbandry

Surgical procedures, experimental protocols and animal care conformed to the NIH Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee at the University of Washington. IACUC Protocol Number 4532-10. Animal husbandry and housing were overseen by the Washington National Primate Research Center. Monkeys had *ad libitum* access to biscuits (Lab Diet, Fiber Plus Monkey Diet Cat#5049) and controlled daily access to fresh produce, treats and water. When possible, animals were pair-housed and allowed grooming contact. Cages were washed every other week, bedding was changed every day, and animals were examined by a veterinarian at least twice per year.

Methods details

10X multiome data generation from mouse and macaque spinal cord

Macaque tissue samples were obtained from the University of Washington National Primate Resource Center. Immediately following euthanasia, macaque spinal cord samples were removed and transported to the Allen Institute in artificial cerebral spinal fluid equilibrated with 95% O_2 and 5% CO_2. Upon arrival at the Allen Institute, spinal cord samples were flash frozen in dry-ice cooled isopentane, transferred to vacuum-sealed bags, and stored at -80°C. To isolate specific regions of interest, spinal cord samples were briefly transferred to -20°C and the region of interest was removed and subdivided into smaller blocks on a custom temperature controlled cold table held at -20°C. After dissection, tissue blocks were returned to -80°C until the time of Multiome processing. Nucleus isolation for 10x Chromium Multiome was conducted as described (dx.doi.org/10.17504/protocols.io.y6rfzd6). Briefly, single nucleus suspensions were incubated with DAPI (4',6-diamidino-2-phenylindole dihydrochloride, ThermoFisher Scientific, D1306) at a concentration of 0.1µg/ml, Alexa Fluor 488 rabbit anti-OLIG2 antibody (Abcam ab225099, 1:2000), and mouse anti-NeuN conjugated to PE (FCMAB317PE, EMD Millipore. 1:500). Controls were incubated with mouse IgG1k-PE Isotype control (BD Biosciences, 555749, 1:250 dilution), Alexa Fluor 488 rabbit IgG Isotype Control (Abcam ab199091, 1:250) or DAPI alone. Single-nucleus sorting was carried out on either a BD FACSAria II SORP or BD FACSAria Fusion instrument (BD Biosciences) using a 130 µm nozzle and BD Diva software v8.0. A standard gating strategy based on DAPI, NeuN, and OLIG2 staining was applied to all samples. Doublet discrimination gates were used to exclude nuclei multiplets. NeuN+, OLIG2+, and OLIG2- nuclei were sorted into separate tubes and were pooled at a targeted ratio of 70% NeuN+, 20% OLIG2-, and 10% OLIG2+ after sorting. Sorted samples were centrifuged and the concentration of nuclei (nuclei/µl) was determined. Nuclei were immediately loaded for 10x Chromium multiome processing following the manufacturer's protocol.

Mice were anaesthetized with 2.5–3% isoflurane and transcardially perfused with cold, pH 7.4 HEPES buffer containing 110 mM NaCl, 10 mM HEPES, 25 mM glucose, 75 mM sucrose, 7.5 mM MgCl2, and 2.5 mM KCl to remove blood (see *Protocols.io* https://doi.org/10.17504/protocols.io.5jyl8peq8g2w/v1). Following perfusion, the spinal cord was dissected quickly, cut into 4 parts (C1-C8, T1-T13, L1-L6, S1-S4 + Co1-Co3), frozen for 2 min in liquid nitrogen vapor and then moved to −80 °C for long term storage following a freezing protocol developed at AIBS (see *Protocols.io* https://doi.org/10.17504/protocols.io.j8nlkodr6v5r/v1).

Nuclei were isolated using the RAISINs method [38] with a few modifications as described in a nuclei isolation protocol developed at AIBS (see *Protocols.io* https://doi.org/10.17504/protocols.io.4r3l22n5pl1y/v1. In short, excised tissue dissectates were transferred to a 12-well plate containing CST extraction buffer. Mechanical dissociation was performed by chopping the dissectate using spring scissors in ice-cold CST buffer for 10 min. The entire volume of the well was then transferred to a 50-ml conical tube while passing through a 100-µm filter and the walls of the tube were washed using ST buffer. Next the suspension was gently transferred to a 15-ml conical tube and centrifuged in a swinging-bucket centrifuge for 5 min at 500 rcf and 4 °C. Following centrifugation, the majority of supernatant was discarded, pellets were resuspended in 100 µl 0.1× lysis buffer and incubated for 2 min on ice. Following addition of 1 ml wash buffer, samples were gently filtered using a 20-µm filter and centrifuged as before. After centrifugation most of the supernatant was discarded, pellets were resuspended in 10 µl chilled nuclei buffer and nuclei were counted to determine the concentration. Nuclei were diluted to a concentration targeting 5,000 nuclei per µl.

For 10x Multiome processing, we used the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Bundle (1000283, 10x Genomics). We followed the manufacturer's instructions for transposition, nucleus capture, barcoding, reverse transcription, cDNA amplification and library construction (see *Protocols.io* https://doi.org/10.17504/protocols.io.bp2l61mqrvqe/v1).

Genome assemblies and annotations

M. musculus (mouse) assembly: mm10, GRCm39 annotation:GCF_000001635.27-RS_2023_04; *M. mulatta* (rhesus monkey) assembly: Mmul_10 (rheMac10), annotation: NCBI Annotation Release 103.

10X Multiome RNA-seq data processing and analysis

Raw sequencing data were processed using cellranger-arc (10x Genomics) to generate single-nucleus RNA-seq (snRNA-seq) UMI count matrices for intronic and exonic reads. We annotated cell classes for mouse and macaque by clustering cells and finding enrichment of existing marker panels for broad spinal cord cell classes [1]. Nuclei were filtered to remove low-quality samples by requiring ≥1000 genes detected per non-neuronal nuclei and requiring ≥2000 genes detected per neuronal nuclei. Counts were normalized to transcripts per million and log transformed using scanpy [39]. Putative doublets were removed by setting a threshold of 0.3 on doublet score. Species correction to align mouse and macaque was performed using scVI. A k-nearest neighbour graph was built using the species corrected latent dimensions from scVI. To visualize cell classes, we performed the uniform manifold approximation and projections (UMAPs) nonlinear dimension reduction technique.

ATAC–seq peak calling and filtering

We used ArchR for snATAC–seq peak calling on pseudo bulk ATAC–seq fragments grouped by cell class per species using the standard ArchR processing pipeline. High quality nuclei were retained based on ATAC-seq metrics: transcription start site enrichment >= 3, fragments > 1000, max fragments < 10,000. Cell class annotations were transferred on the snATAC-seq from the snRNA-seq annotations based on distinct marker gene expression. Fragment files were aggregated into cell class bigwig files using ArchR and uploaded to the UCSC genome browser for identification of putative enhancer elements.

Accessibility screening in external data

Accessibility counter screening in human whole body and mouse whole body was performed by downloading or constructing bigwig files from [21,22]. To assess the accessibility of each putative spinal cord enhancer in these previous studies we used `multiBigwigSummary` from deeptools. This tool produces for each enhancer the average number of fragments within an enhancers genomic coordinate for each ATAC-seq study.

Viral vector construction

Putative enhancers were cloned from C57BL/6J mouse genomic DNA using enhancer-specific primers and Q5 high-fidelity DNA polymerase (NEB Cat#M0494S). Individual enhancer sequences were initially cloned into a rAAV backbone that contained a minimal beta-globin promoter, an SYFP2 (a yellow fluorescent protein), a bovine growth hormone polyA (BGHpA), and a woodchuck post-transcriptional regulatory element (WPRE3) using standard restriction digest cloning or Gibson assembly techniques. The 3xcore of the 1139m enhancer and the

3XFlag-mTFP1 transgenes were synthesized by Genscript and inserted into the standard rAAV backbone using standard restriction digest cloning. All plasmid sequences were verified via Sanger sequencing and restriction digests were performed to confirm intact inverted terminal repeat (ITR) sites.

Viral packaging and titering

Crude and purified rAAVs of the PHP.eB, 9P31, or PHP.S serotypes were generated in house using previously described methods [16]. The average titer of these viral preps was 5.0 X 10^{13} GC/ml.

Mouse retro-orbital (RO) and intracerebroventricular injections (ICV) of AAVs, histology and imaging

C57BL/6J or double transgenic mice were retro-orbitally injected in 21–45-day-old age range (P21-P45). Mice were briefly anesthetized with isoflurane and 5.0 x 10^{11} viral genome copies (GC) were delivered into the right venous sinus in a volume of 90µL or less. 0.5% proparacaine hydrochloride ophthalmic solution (Patterson Cat#07-885-9765) was then applied to the injected eye and animals recovered the same day. Mice were euthanized 3-4 weeks post-infection for analysis.

For ICV injections, C57BL/6J mouse neonatal pups (P0-P3) were injected with 1.0 x 10^{11} GC into the right cerebral ventricle and euthanized 3-4 weeks post-infection for analysis.

On the day of sacrifice, mice were anesthetized with isoflurane and transcardially perfused with 1x phosphate buffered saline (PBS) followed by 4% paraformaldehyde (PFA). Organs of interest were removed, post-fixed in 4% PFA overnight at 4°C, followed by an additional incubation for 2-3 days in 30% sucrose at 4°C. Transverse spinal cord sections or sagittal brain sections or varying angles for peripheral tissue sections were cut using a freezing microtome (Leica SM2000R) or cryostat (Leica CM3050S). Sections thickness was 25-50 µm across the various tissues and sections were stored in 1xPBS containing 0.01% sodium azide (Millipore Sigma Cat#S2002) at 4°C until downstream applications. To visualize direct SYFP2 or tdTomato fluorescence, sections were mounted onto Superfrost slides (Fisher Cat#12-550-15) using VECTASHIELD Hardset antifade mounting medium with DAPI (Vector Laboratories Cat#H-1500-10) and imaged on a Nikon Eclipse Ti epifluorescence microscope or Leica SP8 confocal microscope.

Rat ICV injections of AAVs, immunohistochemistry, and imaging

Rat neonatal pups P0-P1 were injected using the following coordinates to hit the lateral ventricle: 1.5 mm lateral to bregma, 1.5 mm posterior to bregma, and 2 mm deep. A total of ~1.50-2.00 X 10^{11} GC was delivered in 5µL total volume per hemisphere, and the needle was held in place in the ventricle for 30 seconds to allow virus diffusion prior to removal. Injected animals were returned to their home cage with the dam and euthanized at P18-P19 for analysis of viral expression.

For SYFP2 and cholinergic positive neuron detection, fixed sections were washed three times in 1xPBS at room temperature with gentle shaking, blocked in 1xPBS containing 5% normal goat serum (Vector Laboratories Cat#S-1000-20), 0.2% Triton X-100 (Millipore Sigma Cat#X100), bovine serum albumin (Millipore Sigma Cat#A9418-5G) for up to 2 hours at room temperature with gentle shaking, and then incubated overnight at 4°C in the primary antibody. The IHC

experiments noted include the use of chicken anti-GFP (Aves labs Cat#GFP-1020) and mouse anti-ChAT (Atlas Cat#AMAb91130). After primary incubation, sections were washed three times in 1xPBS and incubated in the same blocking solution containing secondary antibodies. Secondary antibodies used include goat anti-chicken Alexa 488 (ThermoFisher Cat#A-11039) and goat anti-mouse IgG2b Alexa 647 (ThermoFisher Cat#A-21247).

Following secondary incubation, sections were washed three times and mounted with ProLong Gold Antifade Mountant (Invitrogen Cat#P36930). Images were taken on a Leica SP8 confocal microscope using the manufacturer's software.

Intra-cisterna magna (ICM) and intrathecal (IT) injections of AAV vectors into macaque, histology, immunochemistry, and imaging

For injections into macaque, large scale purified viral vector preps were packaged of the PHP.eB serotype by a commercial source (Packgene, Houston TX). The average titer of these viral preps was 9.0 X 10^{13} GC/ml. One week prior to survival surgery, an MRI scan was obtained for each animal and used to establish the injection coordinates.

For the IT injection, a three-year-old male macaque weighing 4.53 kg was anesthetized, placed on their side with their hind limbs towards the umbilicus such that there was a widening of the intervertebral space. A 22-gauge spinal needle was inserted into the L3-L4 intervertebral space, and its position was confirmed by CSF return. The syringe containing the virus was attached to the spinal needle and virus was injected using a syringe pump (Medfusion 2001) at a speed of 0.03mL per minute, with 2.27mL total volume injected at 1.0 X 10^{13} vg/kg. The tubing attached to the syringe was backfilled with saline and flushed immediately after to account for the dead space within the tubing.

For the ICM injections, two to three-year-old male and female macaques weighing 3-4 kg were anesthetized and placed in ventral recumbency with their head ventroflexed. A 22-gauge spinal needle was inserted into the area between the occipital protuberance and the C1 vertebrae and its position was confirmed by CSF return. The syringe containing the virus was attached to the spinal needle and virus was injected using a syringe pump (Medfusion 2001) at a speed of 0.03mL per minute, with 2.38mL total volume injected at 1.0 X 10^{13} vg/kg. The tubing attached to the syringe was backfilled with saline and flushed immediately after to account for the dead space within the tubing.

Animals were euthanized 5-7 weeks post-injection and the spinal cord, brain, and several other peripheral organs were collected. Fresh tissue was drop-fixed in 10% formalin for 48 hours at 4°C and then transferred to a 30% sucrose solution and kept at 4°C for several days to cryoprotect. and sub sectioning. Transverse spinal cord sections of 30 µm from each level (cervical, thoracic, lumbar, and sacral) and from brain and other peripheral organs were cut using a freezing microtome (Leica SM2000R) or cryostat (Leica CM3050S).

For Flag-mTFP1 and cholinergic positive neuron detection, fixed sections were washed three times in 1xPBS at room temperature with gentle shaking, blocked in 1xPBS containing 5% normal goat serum (Vector Laboratories Cat#S-1000-20), 0.2% Triton X-100 (Millipore Sigma Cat#X100), bovine serum albumin (Millipore Sigma Cat#A9418-5G) for up to 2 hours at room temperature with gentle shaking, and then incubated overnight at 4°C in the primary antibody. The IHC experiments noted include the use of chicken anti-GFP (Aves labs Cat#GFP-1020), mouse anti-FLAG M2 (Sigma Cat#F1804) and mouse anti-ChAT (Atlas Cat#AMAb91130). After

primary incubation, sections washed three times in 1xPBS and incubated in the same blocking solution containing secondary antibodies. Secondary antibodies used include goat anti-chicken Alexa 488 (ThermoFisher Cat#A-11039), goat anti-mouse IgG1 Alexa 488 (ThermoFisher Cat#A-21121), and goat anti-mouse IgG2b Alexa 555 (ThermoFisher Cat#A-21147).

Following secondary incubation, sections were washed three times and mounted with Vectashield Antifade Mounting Medium with DAPI (Vector Labs Cat#H-1200-10). Images were acquired on a Leica SP8 confocal microscope a a Nikon Eclipse Ti2 epifluorescence microscope using the manufacturer's software.

RNAscope on mouse spinal cord sections and analysis

We performed RO injections of wild-type C57BL/6J mice with enhancer AAVs driving SYFP2, sacrificed animals 4-5 weeks post-injection and harvested the spinal cord tissue. Following dissection, spinal cords were embedded in optimum cutting temperature compound (OCT; TissueTek Cat#4583) and then stored at -80°C until ready to section. Immediately prior to sectioning, the spinal cords were thawed for one hour at –20°C and then cut at a thickness of 20 µm on a cryostat. Thin sections were collected directly onto SuperFrost slides (ThermoFisher Scientific Cat#J3800AMNZ) and then left to dry for up to one hour at -20°C to ensure adhesion of the section to the slide. Slides were kept at –80°C for up to two weeks before single molecule fluorescent *in situ* hybridization was performed using the RNAscope Fluorescent Multiplex Reagent Kit V2 (Advanced Cell Diagnostics Cat#323100). The assay was carried out using the manufacturer's instructions for fresh frozen tissue apart from one small deviation which was to fix the tissue for only 15 minutes in 4% PFA after removing slides from freezer. Probes used were *Chat* (ACD Cat#408731-C2) to label LMNs broadly, *Chodl* (ACD Cat#450211-C3), and *Ret* (ACD Cat#431791-C3) to differentiate between LMN subtypes. Since native SYFP2 fluorescence degrades in fresh frozen sections, a probe against *SYFP2* mRNA (ACD Cat#590291) was used to identify cells labeled by the specified enhancer viruses. All sections were stained with DAPI and mounted with ProLong Gold Antifade mounting medium (Invitrogen Cat#P36930) per the RNAscope kit protocol.

Sections were imaged using a 10x objective on a Leica SP8 confocal microscope with resonance scanning. Approximately 20 µm stacks with 1µm steps were collected and processed in LASX (Leica) software to produce maximum intensity projections for each channel.

Clearing and light sheet imaging of mouse spinal cord

Paraformaldehyde-fixed samples were preserved using SHIELD reagents (LifeCanvas Technologies) using the manufacturer's instructions [40]. Samples were delipidated using LifeCanvas Technologies Clear+ delipidation reagents. Following delipidation, samples were incubated in 50% EasyIndex (RI = 1.52, LifeCanvas Technologies) overnight at 37°C followed by 1 day incubation in 100% EasyIndex for refractive index matching. After index matching, the samples were imaged using a SmartSPIM axially swept light sheet microscope using a 3.6x objective (0.2 NA) (LifeCanvas Technologies).

Quantification and statistical analysis

Cell counts of epifluorescence or confocal images were obtained using the ImageJ analysis Cell Counter plug in. Briefly, the multichannel tiff files were uploaded into ImageJ, split into the individual channels, and brightness and contrast were adjusted uniformly across the images such that the signal could be easily distinguished from the background. The individual channel

images were then compressed into a stack and the number of cells visible in each channel were counted.

For IHC and double transgenic experiments, percent specificity for a given virus was calculated by taking the number of double positive cells (i.e., SYFP2+ and Chat+ or tdTomato+) and dividing that number by the total number of virally labeled cells (i.e. SYFP2+). Percent completeness of labeling was calculated by taking the number of double positive cells (i.e., SYFP2+ and Chat+ or tdTomato+) and dividing that by the total number of positive cells for the cell type of interest population (i.e., Chat+ or tdTomato+).

For RNAscope experiments, we quantified specificity and completeness of our viruses against using subtype selective probes, *Ret* and *Chodl* and the pan LMN marker, *Chat*. Percent specificity for a given virus was calculated by taking the number of double positive cells (i.e., *SYFP2*+ and *Ret*+ or *Chodl*+) and dividing that number by the total number of virally labeled cells (i.e., *SYFP2*+). Percent completeness of labeling was calculated by taking the number of double positive cells (i.e., *SYFP2*+ and *Ret*+ or *Chodl*+) and dividing that by the total number of positive cells for the cell type of interest population (i.e., *Ret*+ or *Chodl*+). Percent specificity for the subtype selective probes was calculated by taking the double positive cells (i.e., *Ret*+ or *Chodl*+ and *Chat*+) and dividing that number by the total number of *Ret*+ or *Chodl*+ cells. Percent completeness of labeling was calculated by taking the number of double positive cells (i.e., *Ret*+ or *Chodl*+ and *Chat*+) and dividing that by the total number of *Chat*+ positive cells.

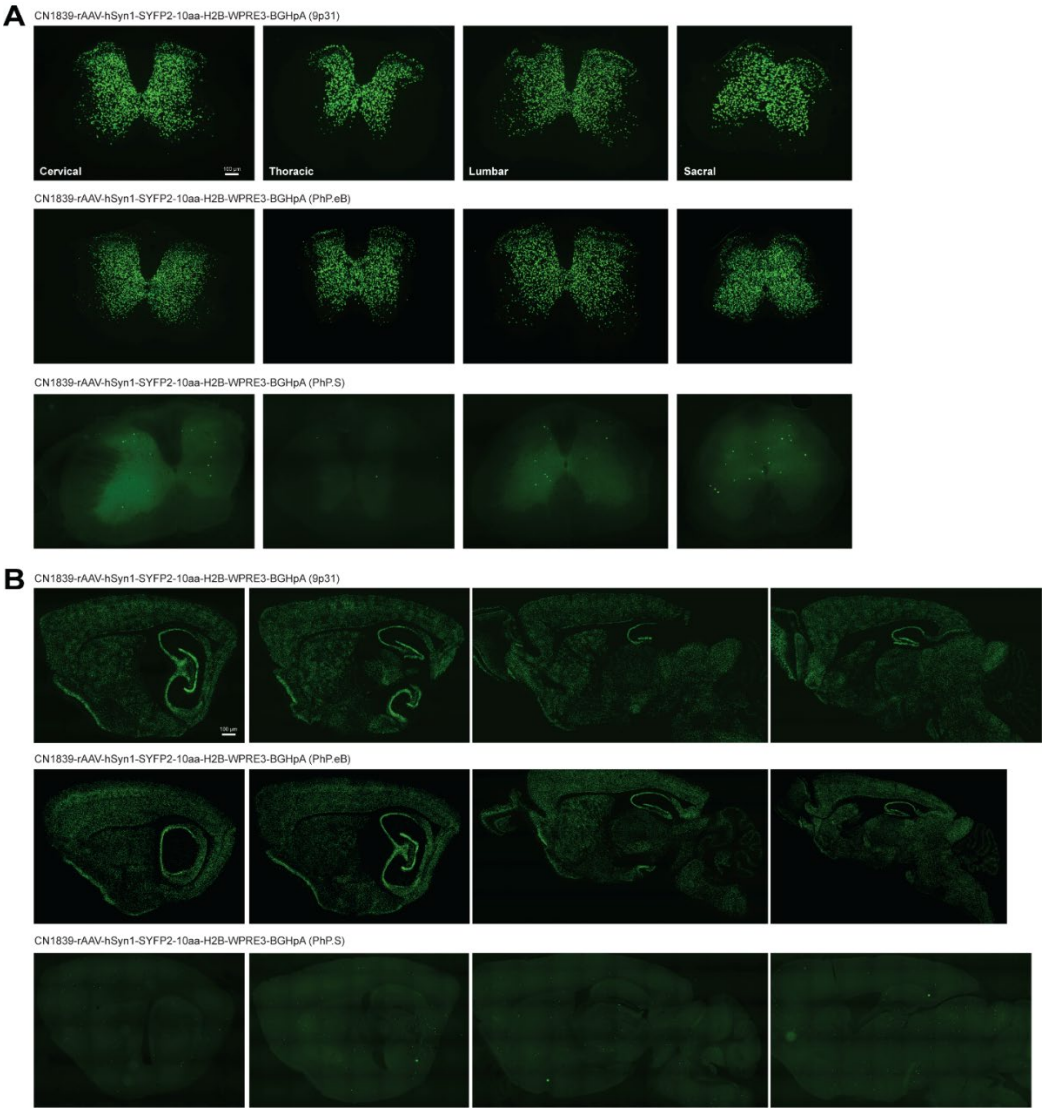
Values were imported, analyzed, and graphed using GraphPad Prism 10 software (Dotmatics, Boston, MA) and data are presented as mean ± standard error from the mean. Statistical comparisons of means within the same treatment group was performed using a one-way ANOVA or using a two-way ANOVA for comparison of the same virus injected via different routes of administration.

Supplemental Information

Document S1. Figures S1-S4 and Table S2.

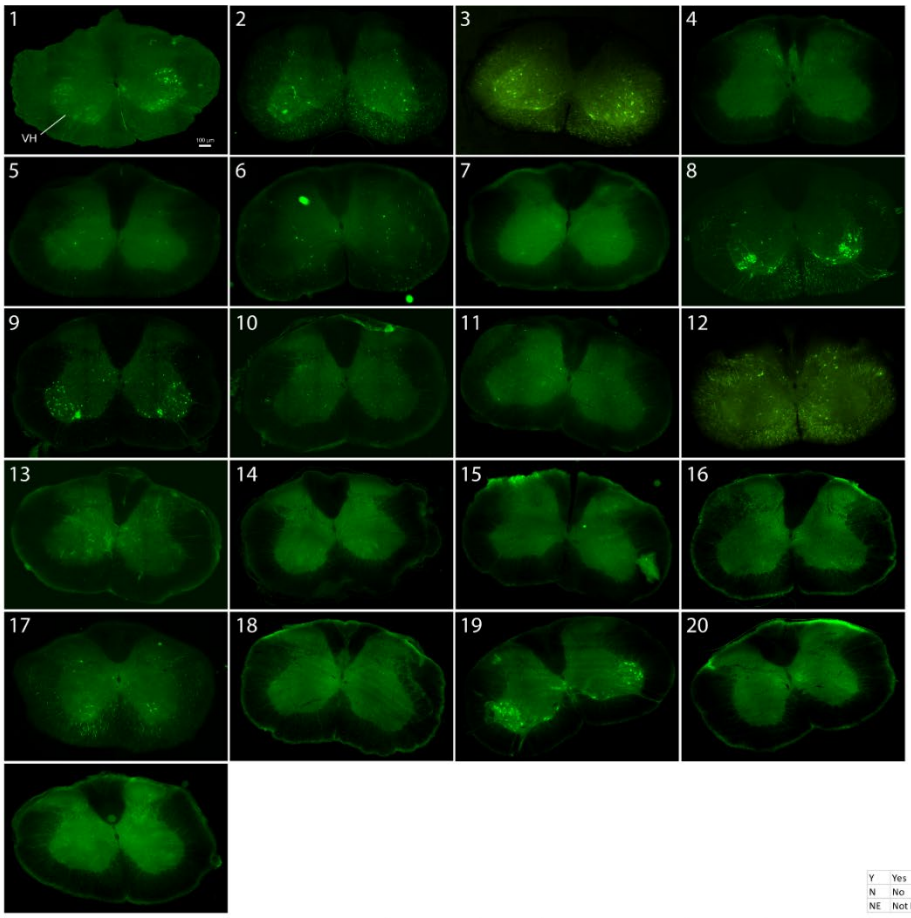
Table S1. Excel file documenting accessibility of pan-LMN enhancers in [21,22], contains additional data too large to fit in a PDF, related to Figure 1E-F

Video S1. Lightsheet video of an HCT1 virus infected mouse spinal cord, related to Figure 2



Supplemental Figure 1. Capsid screening in mouse spinal cord and brain. (**A**) AAVs containing the pan neuronal hSyn1 promoter driving SYFP2 fluorescent protein expression were packaged with three candidate BBB permeant capsids (PHP.eB, 9P31, PHP.S) and tested by RO injection at a dose of 5.0 x 10^{11} GC in wild-type mice. Representative epifluorescence images of native SYFP2 fluorescence in transverse sections from each level of cord are shown for all viruses screened. (**B**) Representative epifluorescence images of native SYFP2 fluorescence in sagittal brain sections prepared from virus infected animals are shown at levels ranging from lateral to medial for all viruses screened.

A

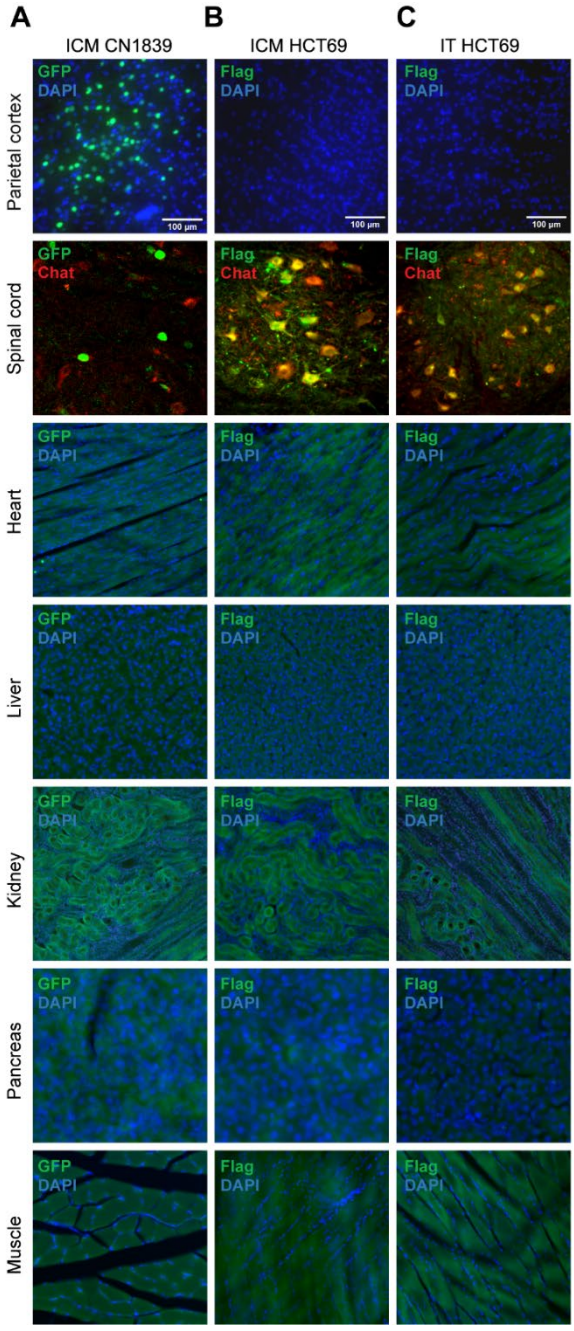


B

Y	Yes
N	No
NE	Not Evaluated

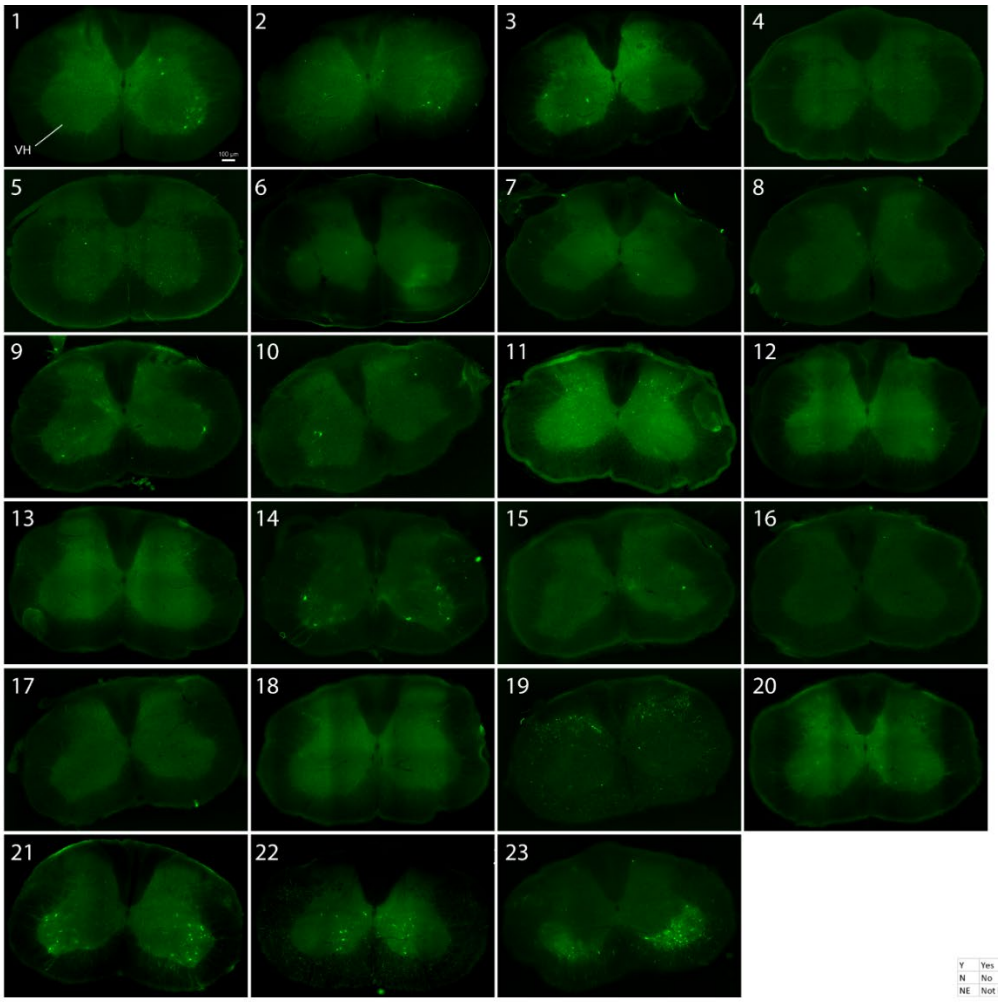
Number	Plasmid #	Full Plasmid Name	Enhancer name	Proximal marker gene	Mouse genome coordinates	SYFP2 expression	SYFP2 expression in ventral horn	Validation (IHC, transgenic and/or RNAscope)
1	HCT1	rAAV-eHGT_1137m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1137m	*ChAT*	chr14:32,468,473-32,469,163	Y	Y	Y
2	HCT2	rAAV-eHGT_1138m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1138m	*ChAT*	chr14:32,265,147-32,265,998	Y	Y	Y
3	HCT6	rAAV-eHGT_1141m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1141m	*Spp1*	chr5:104,239,701-104,240,770	Y	Y	Y
4	HCT7	rAAV-eHGT_1142m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1142m	*Spp1*	chr5:104,470,842-104,471,492	Y	N	NE
5	HCT8	rAAV-eHGT_1143m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1143m	*Esrrg*	chr1:187,052,314-187,053,069	Y	Y	NE
6	HCT9	rAAV-eHGT_1144m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1144m	*Esrrg*	chr1:187,052,266-187,053,125	Y	Y	NE
7	HCT10	rAAV-eHGT_1145m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1145m	*ChAT*	chr14:32,462,648-32,464,201	N	N	NE
8	HCT11	rAAV-eHGT_1048m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1048m	*ChAT*	chr14:32468423-32469127	Y	Y	Y
9	HCT12	rAAV-eHGT_1049m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1049m	*Parg*	chr14:32,265,287-32,265,964	Y	Y	Y
10	HCT13	rAAV-eHGT_1050m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1050m	*Chat*	chr14:32,467,171-32,468,049	Y	Y	NE
11	HCT14	rAAV-eHGT_1051m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1051m	*Ogdh1*	chr14:32,323,065-32,323,649	Y	Y	NE
12	HCT15	rAAV-eHGT_1052m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1052m	*Parg*	chr14:32,226,503-32,227,055	Y	Y	NE
13	HCT19	rAAV-eHGT_1056m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1056m	*Poln*	chr5:34,117,005-34,117,651	N	N	NE
14	HCT33	rAAV-eHGT_1159m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1159m	*Ecel1*	chr1:87,316,608-87,317,271	N	N	NE
15	HCT 75	rAAV-eHGT_1302m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1302m	*Lsl1*	chr13:116,401,153-116,401,863	N	N	NE
16	HCT 76	rAAV-eHGT_1303m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1303m	*Lsl1*	chr13:116,534,265-116,535,361	Y	Y	NE
17	HCT77	rAAV-eHGT_1304m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1304m	*Lsl1*	chr13:116,293,303-116,294,276	Y	Y	NE
18	HCT79	rAAV-eHGT_1306m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1306m	*Nefh*	chr11:4,816,376-4,817,039	Y	N	NE
19	HCT80	rAAV-eHGT_1307m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1307m	*Olig3*	chr10:19,297,683-19,299,746	Y	Y	Y
20	HCT81	rAAV-eHGT_1308m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1308m	*Prph*	chr15:99,044,641-99,045,405	Y	N	NE
21	HCT82	rAAV-eHGT_1309m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1309m	*Prph*	chr15:99,041,013-99,041,841	N	N	NE

Supplemental Figure 2. Initial SYFP2 screening of putative pan-spinal motor neuron enhancer viruses in mouse, related to Figure 2. (**A**) Candidate enhancers were initially screened by RO injection of SYFP2-expressing viruses. Representative epifluorescence images of native SYFP2 fluorescence in transverse cervical sections are shown for all viruses screened. VH = ventral horn and number corresponds to the summary table in panel **B**. Enhancer viruses with no visible cells labelled throughout the sections analyzed were put into “No” SYFP2 expression category, and those viruses that showed some positive SYFP2 cells in any location of the section were put into “Yes” SYFP2 expression category. Viruses labeling cells within the ventral horn were also categorized into “Yes” SYFP2 expression in ventral horn. Promising viruses further validated are noted in the last column.



Supplemental Figure 3. CNS and peripheral organ HCT69 and CN1839 virus expression following ICM or IT routes of administration to macaque. Representative epifluorescence images of anti-Flag signal (green) or anti-GFP signal (green) merged with anti-Chat signal (red) in transverse spinal cord sections prepared from CN1839 ICM (**A**) HCT69 ICM (**B**) or HCT69 IT (**C**) PHP.eb virus injected animals. Representative epifluorescence images of anti-Flag signal (green) or anti-GFP signal (green) merged with anti-DAPI signal (blue) in coronal brain and transverse peripheral tissue sections prepared from CN1839 ICM (**A**) HCT69 ICM (**B**) or HCT69 IT (**C**) PHP.eb virus injected animals.

A



Y	Yes
N	No
NE	Not Evaluated

B

Number	Plasmid #	Plasmid Name	Enhancer name	Target cell subclass/type	Proximal marker gene	Mouse genome coordinates	SYFP2 expression	SYFP2 expression ventral horn	Validation (IHC, transgenic or RNAscope)
1	HCT3	rAAV-eHGT_1139m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1139m	Alpha SMN	*Chodl*	chr16:78,788,098-78,789,110	Y	Y	Y
2	HCT4	rAAV-eHGT_1140m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1140m	Alpha SMN	*Chodl*	chr16:79,718,084-79,718,710	Y	Y	Y
3	HCT39	rAAV-eHGT_1181m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1181m	Alpha SMN	*Calcb*	chr7:114,822,673-114,823,696	Y	Y	NE
4	HCT40	rAAV-eHGT_1182m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1182m	Alpha SMN	*Calcb*	chr7:114,835,534-114,836,229	N	N	NE
5	HCT41	rAAV-eHGT_1183m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1183m	Alpha SMN	*Calcb*	chr7:114,724,500-114,725,335	Y	N	NE
6	HCT42	rAAV-eHGT_1184m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1184m	Alpha SMN	*Calcb*	chr7:114,645,111-114,645,656	Y	Y	NE
7	HCT43	rAAV-eHGT_1185m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1185m	Alpha SMN	*Calcb*	chr7:114,571,685-114,572,623	Y	Y	NE
8	HCT128	rAAV-eHGT_1495m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1495m	Alpha SMN	*Sv2b*	chr7:75,424,131-75,424,880	N	N	NE
9	HCT129	rAAV-eHGT_1496m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1496m	Alpha SMN	*Sv2b*	chr7:75,426,992-75,427,731	Y	Y	NE
10	HCT130	rAAV-eHGT_1497m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1497m	Alpha SMN	*Sv2b*	chr7:75,633,814-75,634,738	Y	Y	NE
11	HCT131	rAAV-eHGT_1498m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1498m	Alpha SMN	*Glis3*	chr19:26,146,054-26,147,089	Y	N	NE
12	HCT132	rAAV-eHGT_1499m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1499m	Alpha SMN	*Glis3*	chr19:26,030,691-26,031,402	Y	Y	NE
13	HCT133	rAAV-eHGT_1500m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1500m	Alpha SMN	*Glis3*	chr19:25,975,005-25,975,662	Y	N	NE
14	HCT134	rAAV-eHGT_1501m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1501m	Alpha SMN	*Glis3*	chr19:29,448,199-29,449,019	Y	Y	NE
15	HCT135	rAAV-eHGT_1502m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1502m	Alpha SMN	*Glis3*	chr19:29,430,295-29,431,605	Y	Y	NE
16	HCT136	rAAV-eHGT_1503m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1503m	Alpha SMN	*Stk32a*	chr18:42,963,393-42,964,294	N	N	NE
17	HCT137	rAAV-eHGT_1504m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1504m	Alpha SMN	*Stk32a*	chr18:42,924,028-42,924,935	N	N	NE
18	HCT138	rAAV-eHGT_1505m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1505m	Alpha SMN	*Stk32a*	chr18:43,961,554-43,962,103	Y	N	NE
19	HCT139	rAAV-eHGT_1506m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1506m	Alpha SMN	*Stk32a*	chr18:43,967,701-43,968,280	Y	N	NE
20	HCT140	rAAV-eHGT_1507m-minBGlobin-SYFP2-WPRE3-BGHpA	eHGT_1507m	Alpha SMN	*Stk32a*	chr18:43,975,238-43,975,795	Y	N	NE
21	HCT44	rAAV-eHGT_1186m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1186m	Gamma SMN	*Ret*	chr6:118,341,476-118,342,612	Y	Y	NE
22	HCT45	rAAV-eHGT_1187m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1187m	Gamma SMN	*Ret*	chr6:118,336,743-118,337,611	Y	Y	NE
23	HCT46	rAAV-eHGT_1188m-minBGprom-SYFP2-WPRE3-bGHpA	eHGT_1188m	Gamma SMN	*Ret*	chr6:117,714,561-117,715,945	Y	Y	Y

Supplemental Figure 4. Initial SYFP2 screening of putative alpha and gamma spinal motor neuron enhancer viruses in mouse, related to Figure 5. (**A)** Representative epifluorescence images of native SYFP2 fluorescence in transverse cervical sections are shown for all viruses screened. VH = ventral horn and number corresponds to the summary table in panel **B**. Virus categorization was performed as described in **Supplement Figure 2** legend.

Table S1. Predicted accessibility of all putative pan LMN enhancers. Accessibility of putative pan LMN enhancers in human whole body ATAC-Seq [21] and whole mouse brain ATAC-Seq [22]. For each enhancer we report the summed accessibility within the enhancer's genomic region for each cell type or tissue.

Table S2. Summary table of brain, spinal cord, and peripheral organ AAV expression. AAV, adeno-associated virus; ICM, intra-cisterna magna; IT, intrathecal. - = no expression, + is 0 to 5 cells, ++ is 6 to 30 cells, +++ is 30 to100 cells. CN1839 is pan neuronal H2B-SYFP2 virus, and HCT69 is 1137m-3XFLAG-mTFP1 virus.

Video S1. Light sheet video of an HCT1 virus infected mouse spinal cord. Native SYFP2 fluorescence (green) is shown throughout intact spinal cord and in transverse spinal cord sections.

References

1. Russ, D.E., Cross, R.B.P., Li, L., Koch, S.C., Matson, K.J.E., Yadav, A., Alkaslasi, M.R., Lee, D.I., Le Pichon, C.E., Menon, V., et al. (2021). A harmonized atlas of mouse spinal cord cell types and their spatial organization. Nat. Commun. *12*, 5722.

2. Blum, J.A., Klemm, S., Shadrach, J.L., Guttenplan, K.A., Nakayama, L., Kathiria, A., Hoang, P.T., Gautier, O., Kaltschmidt, J.A., Greenleaf, W.J., et al. (2021). Single-cell transcriptomic analysis of the adult mouse spinal cord reveals molecular diversity of autonomic and skeletal motor neurons. Nat. Neurosci. *24*, 572–583.

3. Osseward, P.J., 2nd, and Pfaff, S.L. (2019). Cell type and circuit modules in the spinal cord. Curr. Opin. Neurobiol. *56*, 175–184.

4. Alkaslasi, M.R., Piccus, Z.E., Hareendran, S., Silberberg, H., Chen, L., Zhang, Y., Petros, T.J., and Le Pichon, C.E. (2021). Single nucleus RNA-sequencing defines unexpected diversity of cholinergic neuron types in the adult mouse spinal cord. Nat. Commun. *12*, 2471.

5. Liau, E.S., Jin, S., Chen, Y.-C., Liu, W.-S., Calon, M., Nedelec, S., Nie, Q., and Chen, J.-A. (2023). Single-cell transcriptomic analysis reveals diversity within mammalian spinal motor neurons. Nat. Commun. *14*, 46.

6. Stifani, N. (2014). Motor neurons and the generation of spinal motor neuron diversity. Front. Cell. Neurosci. *8*, 293.

7. Lemon, R.N. (2008). Descending pathways in motor control. Annu. Rev. Neurosci. *31*, 195–218.

8. Ragagnin, A.M.G., Shadfar, S., Vidal, M., Jamali, M.S., and Atkin, J.D. (2019). Motor neuron susceptibility in ALS/FTD. Front. Neurosci. *13*, 532.

9. Powis, R.A., and Gillingwater, T.H. (2016). Selective loss of alpha motor neurons with sparing of gamma motor neurons and spinal cord cholinergic neurons in a mouse model of spinal muscular atrophy. J. Anat. *228*, 443–451.

10. Rossi, J., Balthasar, N., Olson, D., Scott, M., Berglund, E., Lee, C.E., Choi, M.J., Lauzon, D., Lowell, B.B., and Elmquist, J.K. (2011). Melanocortin-4 receptors expressed by cholinergic neurons regulate energy balance and glucose homeostasis. Cell Metab. *13*, 195–204.

11. Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T.M., and Sockanathan, S. (1999). Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. Neuron *23*, 659–674.

12. Norante, R.P., Massimino, M.L., Lorenzon, P., De Mario, A., Peggion, C., Vicario, M., Albiero, M., Sorgato, M.C., Lopreiato, R., and Bertoli, A. (2017). Generation and validation of novel adeno-associated viral vectors for the analysis of Ca2+ homeostasis in motor neurons. Sci. Rep. *7*, 6521.

13. Nakano, T., Windrem, M., Zappavigna, V., and Goldman, S.A. (2005). Identification of a conserved 125 base-pair Hb9 enhancer that specifies gene expression to spinal motor neurons. Dev. Biol. *283*, 474–485.

14. Lukashchuk, V., Lewis, K.E., Coldicott, I., Grierson, A.J., and Azzouz, M. (2016). AAV9-mediated central nervous system-targeted gene delivery via cisterna magna route in mice. Mol. Ther. Methods Clin. Dev. *3*, 15055.

15. Mich, J.K., Graybuck, L.T., Hess, E.E., Mahoney, J.T., Kojima, Y., Ding, Y., Somasundaram, S., Miller, J.A., Kalmbach, B.E., Radaelli, C., et al. (2021). Functional enhancer elements drive subclass-selective expression from mouse to primate neocortex. Cell Rep. *34*, 108754.

16. Graybuck, L.T., Daigle, T.L., Sedeño-Cortés, A.E., Walker, M., Kalmbach, B., Lenz, G.H., Morin, E., Nguyen, T.N., Garren, E., Bendrick, J.L., et al. (2021). Enhancer viruses for combinatorial cell-subclass-specific labeling. Neuron *109*, 1449-1464.e13.

17. Mich, J.K., Sunil, S., Johansen, N., Martinez, R.A., Leytze, M., Gore, B.B., Mahoney, J.T., Ben-Simon, Y., Bishaw, Y., Brouner, K., et al. (2023). Enhancer-AAVs allow genetic access to oligodendrocytes and diverse populations of astrocytes across species. bioRxivorg. 10.1101/2023.09.20.558718.

18. Ben-Simon, Y., Hooper, M., Narayan, S., Daigle, T., Dwivedi, D., Way, S.W., Oster, A., Stafford, D.A., Mich, J.K., Taormina, M.J., et al. (2024). A suite of enhancer AAVs and transgenic mouse lines for genetic access to cortical cell types. bioRxivorg. 10.1101/2024.06.10.597244.

19. Hitz, B.C., Lee, J.-W., Jolanki, O., Kagda, M.S., Graham, K., Sud, P., Gabdank, I., Strattan, J.S., Sloan, C.A., Dreszer, T., et al. (2023). The ENCODE Uniform Analysis Pipelines. Res Sq. 10.21203/rs.3.rs-3111932/v1.

20. Lopez, R., Regier, J., Cole, M.B., Jordan, M.I., and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. Nat. Methods *15*, 1053–1058.

21. Zhang, K., Hocker, J.D., Miller, M., Hou, X., Chiou, J., Poirion, O.B., Qiu, Y., Li, Y.E., Gaulton, K.J., Wang, A., et al. (2021). A single-cell atlas of chromatin accessibility in the human genome. Cell *184*, 5985-6001.e19.

22. Zu, S., Li, Y.E., Wang, K., Armand, E.J., Mamde, S., Amaral, M.L., Wang, Y., Chu, A., Xie, Y., Miller, M., et al. (2023). Single-cell analysis of chromatin accessibility in the adult mouse brain. Nature *624*, 378–389.

23. Chan, K.Y., Jang, M.J., Yoo, B.B., Greenbaum, A., Ravi, N., Wu, W.-L., Sánchez-Guardado, L., Lois, C., Mazmanian, S.K., Deverman, B.E., et al. (2017). Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nat. Neurosci. *20*, 1172–1179.

24. Nonnenmacher, M., Wang, W., Child, M.A., Ren, X.-Q., Huang, C., Ren, A.Z., Tocci, J., Chen, Q., Bittner, K., Tyson, K., et al. (2021). Rapid evolution of blood-brain-barrier-penetrating AAV capsids by RNA-driven biopanning. Mol. Ther. Methods Clin. Dev. *20*, 366–378.

25. Samaranch, L., Salegio, E.A., San Sebastian, W., Kells, A.P., Foust, K.D., Bringas, J.R., Lamarre, C., Forsayeth, J., Kaspar, B.K., and Bankiewicz, K.S. (2012). Adeno-associated virus serotype 9 transduction in the central nervous system of nonhuman primates. Hum. Gene Ther. *23*, 382–389.

26. Gray, S.J., Matagne, V., Bachaboina, L., Yadav, S., Ojeda, S.R., and Samulski, R.J. (2011). Preclinical differences of intravascular AAV9 delivery to neurons and glia: a comparative study of adult mice and nonhuman primates. Mol. Ther. *19*, 1058–1069.

27. Kanning, K.C., Kaplan, A., and Henderson, C.E. (2010). Motor neuron diversity in development and disease. Annu. Rev. Neurosci. *33*, 409–440.

28. Mercuri, E., Sumner, C.J., Muntoni, F., Darras, B.T., and Finkel, R.S. (2022). Spinal muscular atrophy. Nat. Rev. Dis. Primers *8*, 52.

29. Andersen, J., Thom, N., Shadrach, J.L., Chen, X., Onesto, M.M., Amin, N.D., Yoon, S.-J., Li, L., Greenleaf, W.J., Müller, F., et al. (2023). Single-cell transcriptomic landscape of the developing human spinal cord. Nat. Neurosci. *26*, 902–914.

30. Bruijn, L.I., Miller, T.M., and Cleveland, D.W. (2004). Unraveling the mechanisms involved in motor neuron degeneration in ALS. Annu. Rev. Neurosci. *27*, 723–749.

31. Patel, T., Hammelman, J., Aziz, S., Jang, S., Closser, M., Michaels, T.L., Blum, J.A., Gifford, D.K., and Wichterle, H. (2022). Transcriptional dynamics of murine motor neuron maturation in vivo and in vitro. Nat. Commun. *13*, 5427.

32. Deng, Q., Wang, S., Huang, Z., Lan, Q., Lai, G., Xu, J., Yuan, Y., Liu, C., Lin, X., Feng, W., et al. (2023). Single-cell chromatin accessibility profiling of cell-state-specific gene regulatory programs during mouse organogenesis. Front. Neurosci. *17*, 1170355.

33. Shu, M., Hong, D., Lin, H., Zhang, J., Luo, Z., Du, Y., Sun, Z., Yin, M., Yin, Y., Liu, L., et al. (2022). Single-cell chromatin accessibility identifies enhancer networks driving gene expression during spinal cord development in mouse. Dev. Cell *57*, 2761-2775.e6.

34. Liu, C., Wang, M., Wei, X., Wu, L., Xu, J., Dai, X., Xia, J., Cheng, M., Yuan, Y., Zhang, P., et al. (2019). An ATAC-seq atlas of chromatin accessibility in mouse tissues. Sci. Data *6*, 65.

35. Kohn, D.B., Chen, Y.Y., and Spencer, M.J. (2023). Successes and challenges in clinical gene therapy. Gene Ther. *30*, 738–746.

36. Lalancette-Hebert, M., Sharma, A., Lyashchenko, A.K., and Shneider, N.A. (2016). Gamma motor neurons survive and exacerbate alpha motor neuron degeneration in ALS. Proc. Natl. Acad. Sci. U. S. A. *113*, E8316–E8325.

37. Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell *174*, 465-480.e22.

38. Drokhlyansky, E., Smillie, C.S., Van Wittenberghe, N., Ericsson, M., Griffin, G.K., Eraslan, G., Dionne, D., Cuoco, M.S., Goder-Reiser, M.N., Sharova, T., et al. (2020). The human and mouse Enteric nervous system at single-cell resolution. Cell *182*, 1606-1622.e23.

39. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15.

40. Park, Y.-G., Sohn, C.H., Chen, R., McCue, M., Yun, D.H., Drummond, G.T., Ku, T., Evans, N.B., Oak, H.C., Trieu, W., et al. (2018). Protection of tissue physicochemical properties using polyfunctional crosslinkers. Nat. Biotechnol. *37*, 73–83.